Science of the Total Environment 667 (2019) 271–286

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Reliability of the global climate models during 1961–1999 in arid and semiarid regions of China

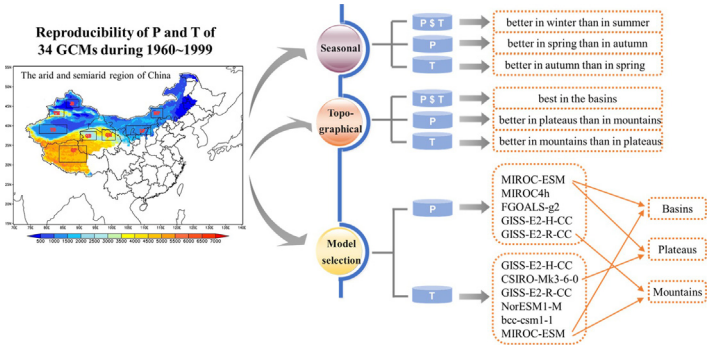
Yanfen Yang [a], Lei Bai [b,c], Bing Wang [a], Jing Wu [d], Suhua Fu [a,*]

[a] State Key Laboratory of Soil Erosion and Dryland Farming on the Loess Plateau, Institute of Soil and Water Conservation, Northwest A&F University, Yangling 712100, Shaanxi, China
[b] College of Resources and Environmental Science, Xinjiang University, Urumqi 830046, China
[c] University of Chinese Academy of Sciences, Beijing 100049, China
[d] Lanzhou Central Meteorological Observatory, Lanzhou 730020, China

HIGHLIGHTS

- The GCMs were evaluated on fine resolution of 0.25° and on three kinds of topography.
- Multiple skill scores and weighting method were used to rank the performance of GCM.
- MIROC-ESM, MIROC4h, FGOALS-g2, GISS-E2-H-CC, GISS-E2-R-CC were the top models for P.
- GISS-E2-H-CC, CSIRO-Mk3-6-0, GISS-E2-R-CC, NorESM1-M, bcc-csm1-1 were the top for T.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 15 October 2018
Received in revised form 8 February 2019
Accepted 12 February 2019
Available online 13 February 2019

Editor: Ouyang Wei

Keywords:
GCM
Precipitation
Temperature
Reproducibility
Seasonal
Topography

ABSTRACT

General circulation models (GCMs) are useful tools for investigating mechanisms of climate change and projecting future climate change scenarios, but have large uncertainties and biases. Accurate models are of significant importance for agriculture, water resources management, hydrological simulation, and species distribution. In this study, we examined the precipitation and temperature reproducibility of 34 GCMs during the period from 1961 to 1999 over arid and semiarid regions of China. The study area was divided into eight sub-regions; each represented a specific topography. The evaluation was conducted for the whole study area and the sub-regions. Spatial and temporal indices and weighting methodology were used to comprehensively illustrate the models' reproducibility. The results showed that the simulation ability during winter outperformed than that during summer (the weight was 0.192 higher for precipitation and 0.044 higher for temperature during winter than that during summer over the whole study area). Precipitation was more accurately simulated during spring than during autumn as opposed to temperature (the weight was 0.124 higher during spring than during autumn for precipitation and 0.1 higher during autumn than during spring for temperature for the whole region). For precipitation, the simulation ability in the basins was the best, followed by plateaus and mountains; the weights were 0.462, 0.308, and 0.231, respectively. For temperature, the mountains and plateaus had the best and poorest reproducibility, at weights of 0.446 and 0.198, respectively. The top models for precipitation and temperature at

* Corresponding author.
E-mail addresses: yfyang@ms.iswc.ac.cn (Y. Yang), suhua@bnu.edu.cn (S. Fu).

https://doi.org/10.1016/j.scitotenv.2019.02.188
0048-9697/

different spatial scales (whole study area, three topography types, eight sub-regions) were recommended. The results served as a reference for model selection in future studies regarding impacts of climate change on eco-hydrology.

1. Introduction

Substantial observational data shows that the global climate has changed over recent decades and changes will likely continue in the future (IPCC, 2007). The Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5) reported that the global average surface temperature has increased by 0.85 °C during the period 1880–2012, with additional warming of 0.3 °C to 0.7 °C expected for the period 2016–2035 relative to the reference period 1986–2005. In addition, precipitation has likely increased since 1901 over the mid-latitude land areas of the Northern Hemisphere (IPCC, 2013).

Arid and semiarid regions account for 45% of the global land area and feed 38% of the world population (Li et al., 2014). In these regions, the environment is fragile because of water shortages with low precipitation, high precipitation variability, and high evaporation (Narisma et al., 2007; Rotenberg and Dan, 2010). Under the influence of global climate change, the responses of energy and water cycle processes are relatively more sensitive in these regions than in other regions (Huang et al., 2012). Arid and semiarid regions contribute approximately 44% to the global annual mean land-surface temperature trend, and have been prone to frequent precipitation events and abrupt changes over the last 100 years (Huang et al., 2012; Reynolds et al., 2007).

The arid and semiarid region of China accounts for 57% of the total land area, mainly in western China. The climate is dry with low precipitation and high evapotranspiration. Because of a soaring population and unreasonable exploitation, many rivers are deprived of water and are drying up (Chen et al., 2011c). The river basins are subject to a severe lack of water resources, resulting in an incongruity between water supply and demand; conflicts between upstream and downstream regions; and a series of ecological and environmental problems in the downstream areas of the desiccated rivers, such as aridification and desertification, groundwater level decreases, desert vegetation degradation, destroyed habitats, and decreased biodiversity (Chen et al., 2006; Feng et al., 2005; Liu and Xia, 2004; Thevs et al., 2015). Furthermore, this vulnerable ecological system is more sensitive to climate change. Precipitation projections based on multiple models predict that this region will be among the most pronounced areas of temperature and precipitation increases in the future (Zhao et al., 2014), significantly affecting the eco-environmental system. Therefore, a better understanding of climate change is of significance for related scientific disciplines.

GCMs are state-of-the-art tools used for investigating mechanisms and causes of climate change and projecting future climate change scenarios (Li et al., 2011). They are also helpful in improving scientific understanding of the earth's climatic system and simulation capabilities in this field (Hirota and Takayabu, 2013). However, they show large biases relative to observational data (Sillmann et al., 2013) because of limited computational resources, incomplete understanding of the manner in which the climate reacts, simplified assumptions in model construction, uncertainties in model parameterization, and unrealistic estimations of natural forcing (Duan and Phillips, 2010; Kiktev et al., 2003; Li et al., 2007; Phillips and Gleckler, 2006). The models also show large uncertainties, and an individual model shows a varying ability to capture climatological variables because of different forcing, the magnitude of the internal variability, and the climatic sensitivity of individual models (Chen and Frauenfeld, 2014; Huang et al., 2013; Wang and Chen, 2014). In addition, it has been shown that no single model is best for all variables or for all regions (Gleckler et al., 2008); no model fell in the best (or worst) five for both temperature and precipitation (Mote and Salathe, 2010). Consequently, the performances of GCMs need to be evaluated before future projections or hydrological application (Park et al., 2016). Testing models' ability to reproduce the "present climate" is important part in evaluating GCM projections (Mote and Salathe, 2010; Su et al., 2013; Walsh et al., 2008).

Several studies have evaluated GCMs in simulating precipitation and temperature over China (Chen and Frauenfeld, 2014; Xu and Xu, 2012; Yue et al., 2016). The results showed that, at a spatio-temporal scale, GCMs can reproduce the seasonal and interannual change features of temperature, as well as the spatial pattern and seasonal variability characteristics of precipitation (Chen and Frauenfeld, 2014; Su et al., 2013; Zhao et al., 2016). The ensembled models performed well in reproducing warming tendencies but were less effective for precipitation (Xu and Xu, 2012; Zhao et al., 2014). In addition, the models underestimated the temperature (1.8 °C) and overestimated the precipitation (263 mm) in most regions, while underestimating summer precipitation over southeastern China (Chen and Frauenfeld, 2014; Yue et al., 2016). The standard deviations (STDEVs) of the simulated temperatures were higher in the western parts than in the eastern parts of China, while higher STDEVs for precipitation occurred in the southern parts (Xu and Xu, 2012). Topographically, the models can capture the geographical distribution of precipitation and temperature (Xu and Xu, 2012). However, the mean absolute errors (MAEs) of temperature and precipitation simulated by the three scenarios of Coupled Model Intercomparison Project Phase 5 (CMIP5) were higher on the Tibetan Plateau, in the arid area, on the Loess Plateau, in the Sichuan Basin, and on the Yunnan-Guizhou Plateau than those in other regions. Among the MAEs, those of the Tibetan Plateau were the largest, followed by the arid area and the Loess Plateau (Yue et al., 2016). Chen and Frauenfeld (2014) also showed that the models overestimated the precipitation along the eastern edge of the Tibetan Plateau to a large extent, particularly during summer. Several studies showed that it is difficult to appropriately reproduce atmospheric processes in complex terrain (Chen et al., 2010; Su et al., 2013; Zhou and Li, 2002) and more uncertainties appeared in regions of rapid elevational change (Bader et al., 2008). The largest biases often appeared in regions with varying topography (Mao and Robock, 1998).

The aforementioned reviews indicated that the previous studies have expended considerable effort in GCM evaluation and achieved valuable conclusions. Nevertheless, most evaluated the models at a large spatial scale, such as the arid area or northwestern and southern parts of China, although a few studies assessed at a relatively smaller scale and in areas of special topography such as the Tibetan Plateau, Loess Plateau, and Inner Mongolian Plateau (Su et al., 2013; Yue et al., 2016). The arid and semiarid area of China is characterized by complex terrain, including three plateaus, three basins, and two mountains. Each type of topography is important and closely relates to the eco-hydrological cycle of runoff generation-confluence-moisture dissipation at a watershed scale. The reproducibility of GCMs may also differ in topography. Taking them as a whole may neglect the detailed performance of GCMs for each unique topography and their differences. Thus, it would be helpful if GCM performance in the eight subregions was distinguished. By doing so, differences in model performance among regions would be evident and the top models for each region could be recommended to facilitate the models' application in eco-hydrological processes at a small watershed scale.

Statistical metrics play a decisive role in model assessment and ranking, because the best models change with different evaluation indicators (Su et al., 2013). Statistical indicators such as correlation coefficient

(CC), root-mean-square error (RMSE), STDEVs, MAEs, and bias are commonly used (Su et al., 2013; Xu and Xu, 2012; Yue et al., 2016). However, the ranking under one indicator would be different to that of the others, making it difficult to determine the best model. In this study, multi-metrics and weighting methodology were used in the model ranking, which is helpful in objectively selecting models and reducing the uncertainties in climatic projections and eco-hydrological application.

In addition, previous studies evaluated the models at a coarse resolution, e.g., 2.5° × 2.5°, 1° × 1° and 0.5° × 0.5° (Chen and Frauenfeld, 2014; Xu and Xu, 2012; Zhao et al., 2014; Zhao et al., 2016). Coarse resolution is generally more applicable at a continental or global scale. At a regional or hydrological basin scale, the assessment may not produce a fine result. Thus, we evaluated the models at a finer resolution of 0.25° × 0.25°, on which the climate series will be more representative.

Summarily, our objective was to investigate the reproducibility of precipitation and temperature of multiple GCMs in eight regions with specific topography over the arid and semiarid region of China, using multiple evaluating metrics. The results may provide insight into the data requirements of climate change models used in eco-hydrological research.

2. Study area, data, and methodologies

2.1. Study area

The study area is in the arid and semiarid region of China which has an annual precipitation of less than 500 mm as shown in Fig. 1 (Zhao et al., 2014). The area covers 5.46 × 10^6 km^2 and accounts for approximately 56.9% of the total area of China. The elevation varies from −146 to 7247 m. Basins, mountains, and plateaus alternate in a staggered manner in the region. Based on the complicated topography, the area has been divided into eight sub-regions (Chen, 2010) as shown in Fig. 1: Jungar Basin (JB), Tianshan Mountains (TM), Tarim Basin (TB), Qaidam Basin (QB), Qilian Mountains (QM), Qinghai-Tibet Plateau (QP), Inner Mongolian Plateau (MP), and Loess Plateau (LP). The scope and area of the eight sub-regions are listed in Table 1. The area of each sub-region was sufficiently large to represent the topography and not overlap with the adjacent sub-regions.

2.2. Data

CMIP5 has publicly released more than 50 GCMs (http://www.ipcc-data.org/sim/gcm_monthly/AR5/Reference-Archive.html). A total of 34 GCM models, containing both historical data and three representative concentration pathway (RCP) scenarios, were examined in this study. The monthly precipitation and temperature during the period 1961–1999 were analyzed based on the ending time of most models in the CMIP5 archive. The models are listed in Table 2. The delta method (Eqs. (1) and (2)) (Anandhi et al., 2011), a bias correction downscaling technique, was used to downscale the climatological precipitation and temperature data to a spatial resolution of a 0.25° × 0.25° grid during the period from 1961 to 1999 as follows:

$$T_{downscaling}(t) = (T_R)(t) + \left(\overline{T_{obs1999-1961}} - \overline{T_{R1999-1961}}\right)_{mon(i)} \quad (1)$$

$$P_{downscaling}(t) = (P_R)(t) \times \left(\frac{\overline{P_{obs1999-1961}}}{\overline{P_{R1999-1961}}}\right)_{mon(i)} \quad (2)$$

where T and P indicate temperature and precipitation, respectively. R is the raw data in the GCMs. “obs” is observational data. “mon” means that

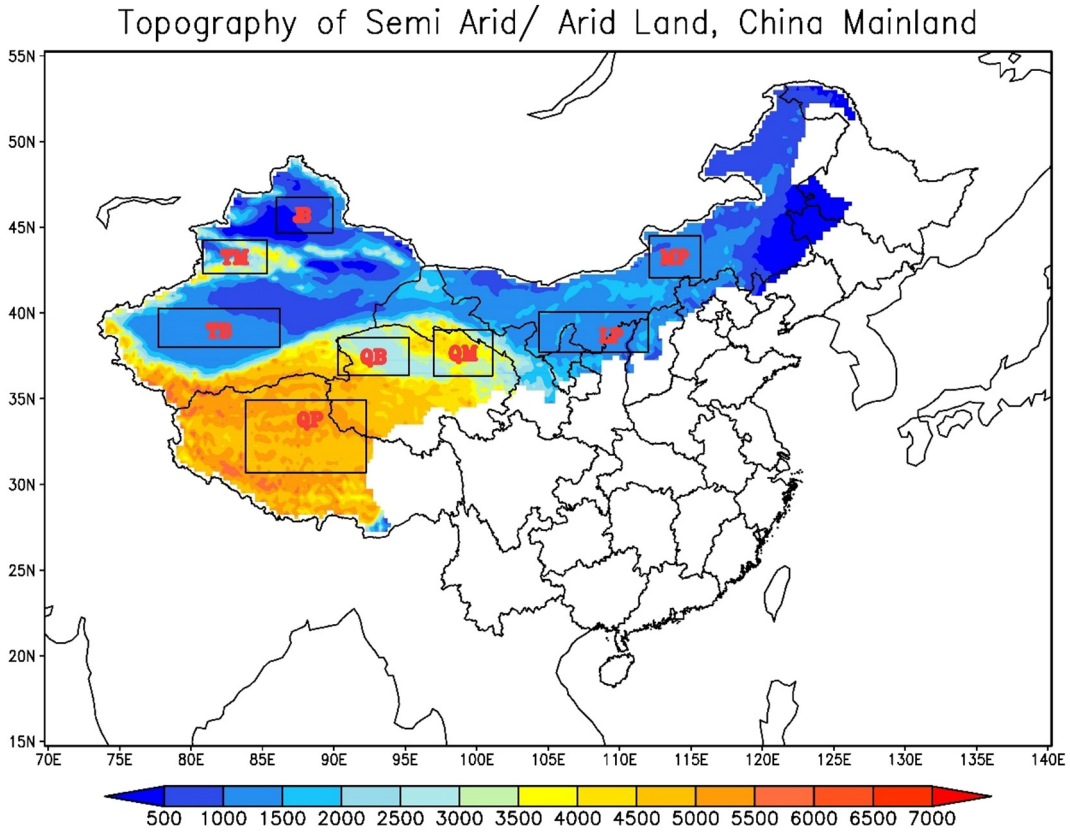


Fig. 1. The location, elevation, province boundary, sub-regions in arid and semiarid region of China. JB (Jungar Basin, 44.67–46.76°N, 85.98–89.92°E), TM (Tianshan Mountains, 42.29–44.25°N, 80.83–85.31°E), TB (Tarim Basin, 37.99–40.25°N, 77.72–86.23°E), QB (Qaidam Basin, 36.36–38.57°N, 90.29–95.25°E), QM (Qilian Mountains, 36.33–39.00°N, 96.98–101.11°E), QP (Qinghai-Tibet Plateau, 30.69–34.91°N, 83.82–92.26°E), MP (Inner Mongolian Plateau, 42.04–44.51°N, 112.06–115.67°E), LP (Loess Plateau, 37.70–40.07°N, 104.36–112.00°E). The elevation unit is m.

Table 1
The scope and areas for 8 sub-regions.

Sub-region	Right lower		Upper left		Area
	Latitude (°)	Longitude (°)	Latitude (°)	Longitude (°)	
JB	44.542	90.217	46.246	85.862	1.704° × 4.355°
TM	42.203	85.558	43.629	81.365	1.426° × 4.192°
TB	37.897	86.000	39.824	77.524	1.927° × 8.476°
QB	35.954	95.257	37.899	89.711	1.945° × 5.546°
QM	35.910	101.629	38.870	96.922	2.960° × 4.707°
QP	29.812	91.194	34.010	82.270	4.198° × 8.924°
MP	42.055	115.274	44.033	112.147	1.978° × 3.127°
LP	37.831	111.667	39.871	103.423	2.040° × 8.244°

calculations are completed by month. The downscaling is the final bias corrected result for temperature and precipitation. "mon(i)" indicates the month for the annual mean monthly precipitation and temperature, mon(i) = 1, …,12. "t" is the month of the calendar year $t = 1, \ldots, 12$. The overline indicates the mean value.

The observed daily precipitation and temperature datasets were obtained from CN05.1 developed by Wu and Gao (2013). The data, with a spatial resolution of 0.25°, were interpolated using the thin plate spline method based on a dataset of 2413 meteorological stations in China from 1961 to 2015. The data from 1961 to 1999 were used in this study.

2.3. Methodologies

The precipitation and temperature of 34 GCMs were examined using the three indices spatial skill score (SS), interannual variability metric (M2), and Taylor skill score (TI). The scores were computed for the entire period 1961–1999; the detailed information is as follows.

2.3.1. Spatial skill score

The SS (Eq. (3)) accounts for the correlation coefficient and standard deviation between the GCM models and ground observations. A model field identical to observations has a skill score of 1. The closer the SS value is to 1, the greater skill in simulating the spatial climatology as follows:

$$SS = 1 - \frac{MSE(m,o)}{MSE(\overline{o},o)} \tag{3}$$

$$MSE = (\overline{m} - \overline{o})^2 + s_m^2 + s_o^2 - 2 s_m s_o r_{m,o} \tag{4}$$

where MSE is the mean squared error and the subscripts m and o are the seasonal GCM models and observations, respectively. Overbars indicate mean; $r_{m,o}$ is the product moment spatial correlation coefficient between the seasonal GCM models and observations; and s_m and s_o are the sample spatial standard deviation of the seasonal model variable and observations, respectively (Murphy, 1988; Pierce et al., 2009).

Table 2
The 34 GCMs evaluated in this study.

No.	Model	Model center	Spatial resolution (latitude × longitude)
1	ACCESS1-0	Commonwealth Scientific and Industrial Research Organization/Bureau of Meteorology, Australia	145 × 192
2	ACCESS1-3		145 × 192
3	bcc-csm1-1	Beijing Climate Center, China	64 × 128
4	bcc-csm1-1-m		160 × 320
5	BNU-ESM	Beijing Normal University, China	64 × 128
6	CanESM2	Canadian Centre for Climate Modelling and Analysis, Canada	64 × 128
7	CCSM4	National Center for Atmospheric Research, USA	192 × 288
8	CESM1-BGC	National Center for Atmospheric Research, USA	192 × 288
9	CESM1-CAM5		192 × 288
10	CESM1-FASTCHEM		192 × 288
11	CESM1-WACCM		96 × 144
12	CSIRO-Mk3-6-0	Commonwealth Scientific and Industrial Research Organization/Queensland Climate Change Centre of Excellence, Australia	96 × 192
13	EC-EARTH	EC-EARTH consortium published at Irish Centre for High-End Computing, Netherlands/Ireland	160 × 320
14	FGOALS-g2	Institute of Atmospheric Physics, Chinese Academy of Sciences, China	60 × 128
15	FIO-ESM	The First Institute of Oceanography, SOA, China	64 × 128
16	GFDL-ESM2G	Geophysical Fluid Dynamics Laboratory, USA	90 × 144
17	GFDL-ESM2M		90 × 144
18	GISS-E2-H-CC	NASA/GISS Goddard Institute for Space Studies, USA	90 × 144
19	GISS-E2-H		90 × 144
20	GISS-E2-R-CC		90 × 144
21	GISS-E2-R		90 × 144
22	HadGEM2-AO	National Institute of Meteorological Research, Korea Meteorological Administration, South Korea	145 × 192
23	HadGEM2-CC	Met Office Hadley Centre, UK	145 × 192
24	HadGEM2-ES		145 × 192
25	INMCM4	Russian Academy of Sciences, Institute of Numerical Mathematics, Russia	120 × 180
26	IPSL-CM5A-LR	Institute Pierre-Simon Laplace, France	96 × 96
27	IPSL-CM5A-MR		143 × 144
28	IPSL-CM5B-LR		96 × 96
29	MIROC4h	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan	320 × 640
30	MIROC5	Agency for Marine-Earth Science and Technology, Japan	128 × 256
31	MIROC-ESM-CHEM		64 × 128
32	MIROC-ESM		64 × 128
33	NorESM1-ME	Bjerknes Centre for Climate Research, Norwegian Meteorological Institute, Norway	96 × 144
34	NorESM1-M		96 × 144

2.3.2. Interannual variability metric

The variability is measured by the interannual standard deviation (STD) and is defined by W. Chen et al. (2011) as follows:

$$M2 = \left(\frac{STD_m}{STD_o} - \frac{STD_o}{STD_m}\right)^2 \quad (5)$$

where STD_m and STD_o denote the interannual standard deviation of the seasonal simulated and observed variables, respectively. M2 is dimensionless and the optimal value is 0; the closer the M2 value is to 0, the greater the ability in simulating the interannual variability.

2.3.3. Taylor skill score

The skill score defined by Taylor (2001) was calculated to examine model skill in reproducing seasonal mean precipitation and temperature. It quantifies the similarity between the distribution and amplitude of the spatial pattern of the model to that of the observations (Hirota et al., 2011).The score is defined as follows:

$$TI = \frac{(1 + r_{m,o})^4}{4(SDR + 1/SDR)^2} \quad (6)$$

where SDR is the ratio of the spatial standard deviation in the models against that of the observations. A higher TI value indicates better simulating skill.

Except for the aforementioned three single indices, the weighting methodology proposed by W. Chen et al. (2011) was used to comprehensively examine the reproducibility. Using this method, all models were first ranked according to their order of performance in terms of each individual index. Then, the weight (W_i) of each model was calculated according to the sum of its rankings for all metrics as follows:

$$R_i = \frac{\sum_{i=1}^{n} S_i}{S_i} \quad (7)$$

$$W_i = \frac{R_i}{\sum_{i=1}^{n} R_i} \quad (8)$$

where S_i is the sum of a model's rank, R_i is the model reliability factor, W_i is the model weight, and n is number of the model. W_i represents the skill of a model in simulating precipitation and temperature. The greater the W_i value, the better the skill; the sum of the weights for all models is equal to 1. Similarly, the eight sub-regions were ordered by the rankings column from best to worst performance during each individual season for every evaluation index, and then the weight of each sub-region was determined (Schuenemann and Cassano, 2009).

3. Results

Based on the aforementioned skill scores, the simulating capabilities of the models in the arid and semiarid region of China were examined. The results are shown in Figs. 2–7; the corresponding values are listed in Table 3.

3.1. Spatial skill score

The SSs for precipitation in the entire study area are −0.52 during spring, −1.07 during summer, −1.03 during autumn, and −0.43 during

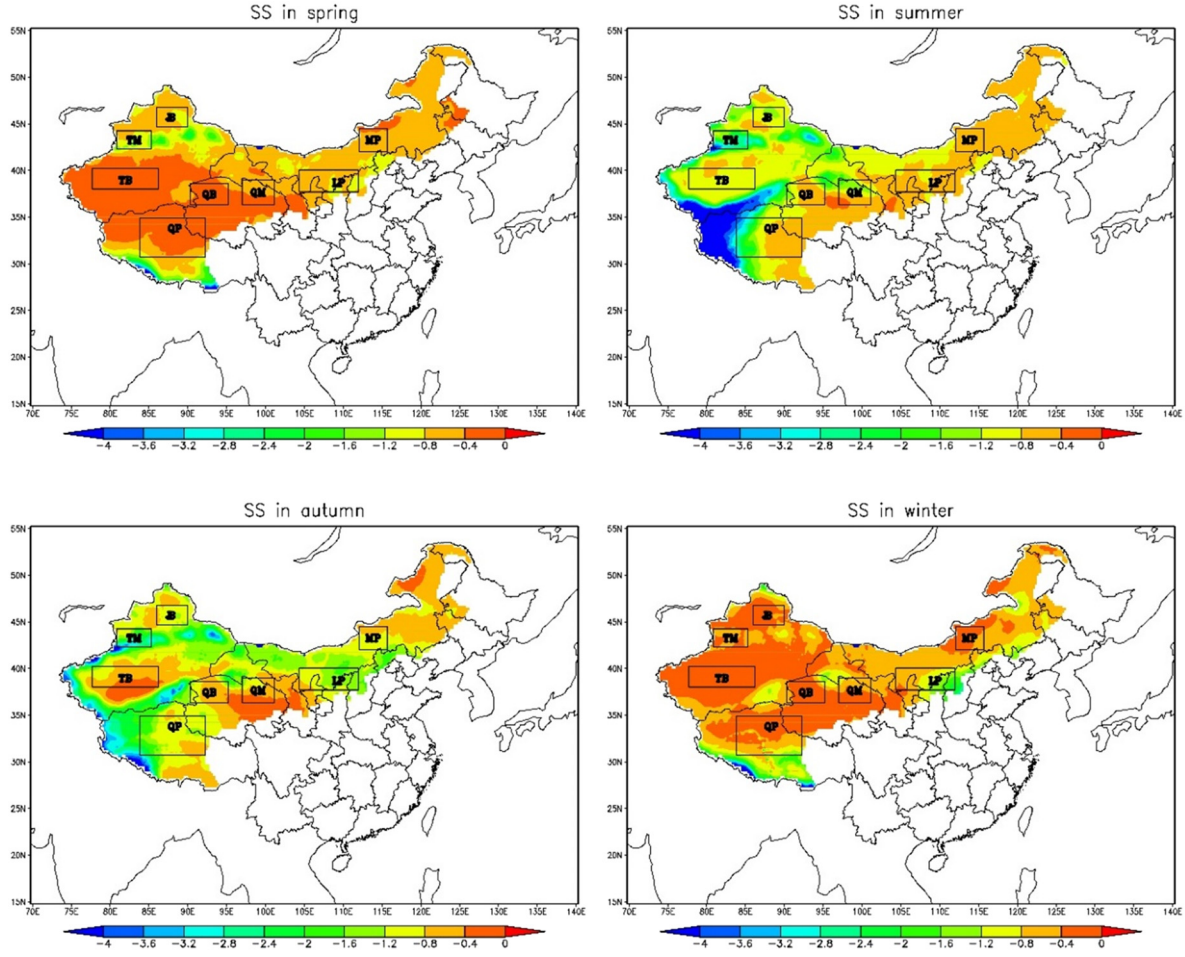


Fig. 2. Spatial skill score (SS) of seasonal precipitation.

winter (Fig. 2, Table 3). The SS varies with season; the model reproducibility during winter and spring are much higher than during summer and autumn. Combining the four seasons, the average SS indicates that the simulating skill of the models differs among regions. For basins, the precipitation reproducibility for the TB (SS = −0.46) and QB (SS = −0.51) are higher than that of the JB (SS = −0.78). For mountains, the SS performance in the QM (−0.63) is much better than that in the TM (−1.44). For plateaus, the SS is best on the MP (−0.6), followed by the QP (−0.79) and LP (−0.9). Overall, the models show the best simulating skill in the basins, followed by plateaus and then mountains. The average SSs are −0.58, −0.76, and − 1.03, respectively. This result indicates that the GCM performance is significantly affected by topography.

The SSs for temperature, by season, are 0.60 (spring), −0.58 (summer), −0.46 (autumn) and −0.44 (winter). This indicates that reproducibility during winter and autumn are better than that during summer and spring (Fig. 3, Table 3). In addition, the temperature seasonal differences are smaller than the precipitation differences. Combining data from all four seasons shows that SS values for temperature differ among the sub-regions, and the differences are not the same as those of precipitation. The models show the best and the poorest results in the JB and TB among the basins, with average SS values of −0.42 and −0.63, respectively. The performance on the LP (SS = −0.5) is higher than that on the QP (SS = −0.69). In addition, the models shown the same simulating skill in the two mountainous areas, with both SS values equal to −0.49. The average simulating skill is best in the mountains (−0.49), followed by the basins (−0.52) and plateaus (−0.54).

3.2. Interannual variability metric

M2 was calculated according to Eq. (5) to investigate model capability in simulating temporal trends of precipitation and temperature. The spatial distributions of M2 for precipitation are shown in Fig. 4 and Table 3. The color bar ranges for the four seasons are not unified because the magnitudes of M2 are distinctly different during the seasons; one color may represent different values for the four subgraphs. The average M2 values for precipitation in the sub-regions are 0.72 (spring), 2.83 (summer), 1.99 (autumn) and 2.09 (winter). This indicates that the consistency of the temporal trend between observations and simulations is best during spring and poorest during summer. For the basins, the best reproducibility appears in the JB, followed by the QB and TB, with M2 values of 1.54, 1.71, and 3.15, respectively. For the plateaus and mountains, the cases are similar to those of the precipitation SSs described in Section 3.1. The models show the best skill on the plateaus, followed by the basins and mountains, with average M2 values of 1.29, 2.14, and 2.48, respectively.

The M2 values shown in Fig. 5 are lower than those of precipitation (Fig. 4), denoting a greater consistency in the temporal trend for temperature between models and observations than that for precipitation. For the entire study area, the M2 of the temperature changes with the season. The simulation skill during autumn is the best (M2 = 0.001), followed by that during spring (M2 = 0.005). The skills during winter (M2 = 0.140) and summer (M2 = 0.149) are much lower. In the sub-regions, the M2 performances during autumn (0.001–0.004) and spring (0.002–0.016) always rank first and second for all sub-regions. The skill during summer ranks third for sub-regions JB, QB, QM and QP, and

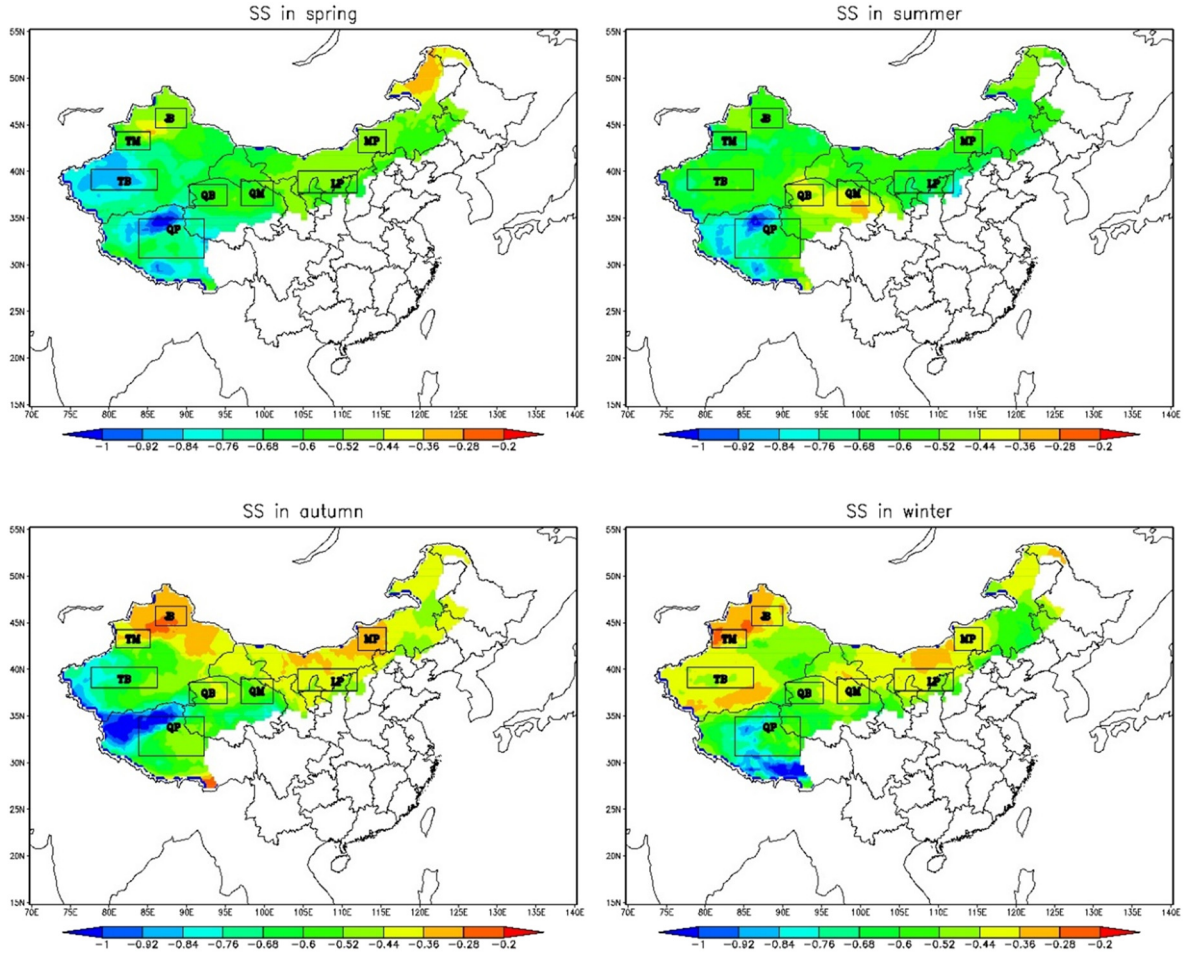


Fig. 3. Spatial skill score (SS) of seasonal temperature.

fourth for the other sub-regions. Ensemble seasonal results show that the models provide the best simulation in the QB for basins, followed by the TB and JB, with M2 values of 0.051, 0.063, and 0.065, respectively. For the plateaus, the skill order is MP, LP, and QP, with M2 values of 0.044, 0.061, and 0.176, respectively. For mountains, the skill in the QM (M2 = 0.043) is much higher than that in the TM (M2 = 0.085). The average M2 for the basins (0.06) and the mountains (0.064) is similar. Reproducibility on the plateaus is much lower, with an M2 value of 0.094.

3.3. Taylor skill score

Distribution of the TI values of seasonal precipitation are shown in Fig. 6 and Table 3. The lowest mean TI of 0.001 during summer indicates the poorest similarity of the distribution and amplitude of the spatial pattern between the models and observations. In contrast, the models provide the highest TI of 0.053 and the best simulation skill during winter. The performance levels during spring (0.024) and autumn (0.024) are moderate and similar. The precipitation reproducibility in the sub-regions mostly follows the aforementioned ranking, with TI ranges of 0.002–0.102 during spring, 0.0004–0.003 during summer, 0.002–0.127 during autumn, and 0.010–0.130 during winter. Taking the four seasons together, the precipitation reproducibility (TI) in the QB (0.075) is much higher than in the JB (0.026) and TB (0.007). The simulating skills on the three plateaus are similar, but slightly higher on the QP. Comparatively, the models perform best in basins with an average TI of 0.036, followed by mountains (TI = 0.022) and plateaus (TI = 0.018).

Comparing Figs. 7 to 6, the model reproducibility for temperature is much better than that of precipitation. For the entire study area, the models provide the best (TI = 0.65) and the poorest (TI = 0.11) temperature simulating skill during autumn and winter, respectively. Performance during spring and summer rank second and third, respectively. This ranking is also applicable for the eight sub-regions. The performances in the QB (0.527) and TB (0.419) are much better than that in the TB (0.253). The TIs on the LP and QP are similar and approximately 0.08 higher than on the MP. The sub-region QM (TI = 0.506) shows better simulating skill than that in the TM (TI = 0.317). Overall, the models have similar performances in the mountains (TI = 0.412) and basins (TI = 0.4) and somewhat higher than on the plateaus (TI = 0.374).

Detailed analysis of the aforementioned three evaluating indices interprets the reproducibility of model precipitation and temperature estimates from different perspectives. The simulating skill reflected by each index is different from that of the other two indices. For example, the models show a higher SS for precipitation on the MP than on the QP, but the opposite is true for TI. Hence, it is difficult to reach a definitive conclusion regarding the performance order regarding the seasons and the sub-regions. By combining information from the three indices, the weights were calculated according to the ranks for each season or topography. This may be a more effective means to determine overall model performance.

3.4. Comprehensive examination of precipitation reproducibility

For the entire study area, the models perform best during winter and spring. The weights are both 0.329 (Table 4) and exceed the mean (0.25). These are followed by weights for autumn (0.205) and summer (0.137). At a sub-region scale, the performances during winter are

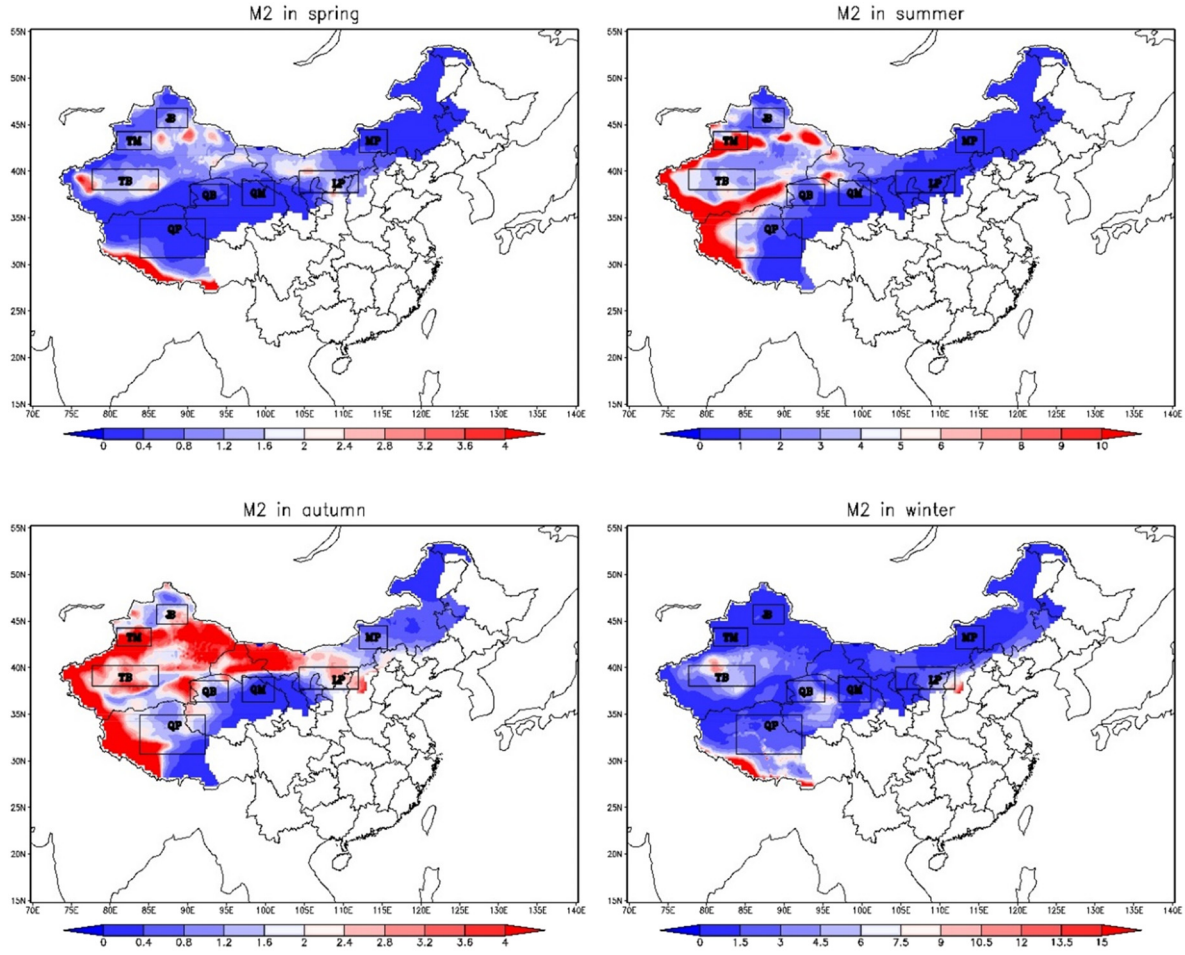


Fig. 4. Interannual variability statistic (M2) of seasonal precipitation.

better than those during summer, except for sub-region LP. The range of the simulating skills during spring (from 0.240 to 0.411) is always better than that during autumn (from 0.160 to 0.269).

The weights for the sub-regions were calculated and are shown in Table 5. The models simulate precipitation better in sub-regions QB, QM, MP, and TB, with the weights all higher than the 0.125 mean. For the three topographic types, the best reproducibility is in the basins (weight = 0.462), followed by plateaus (0.308) and mountains (0.231). For the mountains, the simulating skill in the QM is better than that in the TM. For the plateaus, the best and the poorest reproducibility are on the MP and LP, respectively. For the basins, the order, from best to the worst, is QB > TB > JB.

The simulating skill of individual models, however, is different among the sub-regions (Table 6). For the entire study area, the top five models and their weights are MIROC-ESM (0.084), MIROC4h (0.078), FGOALS-g2 (0.051), GISS-E2-H-CC (0.051) and GISS-E2-R-CC (0.049). The results recommend use of MIROC-ESM for the basins and plateaus, and GISS-E2-R-CC for the mountains. The recommended top five models for each region are listed in Table 6; the weight values are marked in bold.

3.5. Comprehensive examination of temperature reproducibility

The simulating capability of temperature is opposite to that of precipitation. For the entire study area and at a sub-region scale, the reproducibility of the temperature during autumn is higher than during spring (Table 4), which is opposite to the precipitation results. For the entire study area and sub-regions TM, MP, and LP, the skill during winter is better than that during summer, but is the opposite for sub-regions TB, QB, QM, and QP.

Similar to precipitation, the best skills, and weights, are in the QM (0.218) and QB (0.153) and on the MP (0.139); the weights in the other sub-regions are all lower than the mean (0.125) (Table 5). Different from precipitation, the models and associated weights are best in the mountains (0.446), followed by the basins (0.356) and plateaus (0.198). For the three topographic types, model temperature performances are similar to those of precipitation, but the differences lie in the reproducibility for the LP which is better than that of the QP.

Individual models show different simulating skill among the sub-regions (Table 7). For the study area, the models (and weights) GISS-E2-H-CC (0.106), CSIRO-Mk3-6-0 (0.083), GISS-E2-R-CC (0.061), NorESM1-M (0.043) and bcc-csm1-1 (0.041) are recommended. MIROC-ESM shows the best skill in the basins and mountains, and CSIRO-Mk3-6-0 is the best on the plateaus, with weights of 0.072, 0.089, and 0.19, respectively. The recommended top five models for each region are also listed in Table 7; the weight values are marked in bold.

4. Discussion

4.1. Reproducibility depends on topography and season

In this study, the precipitation and temperature reproducibility of 34 GCMs during the period 1961–1999 over an arid and semiarid area of China were examined. Model performance was influenced by season and topography. Seasonally, the models show better performance

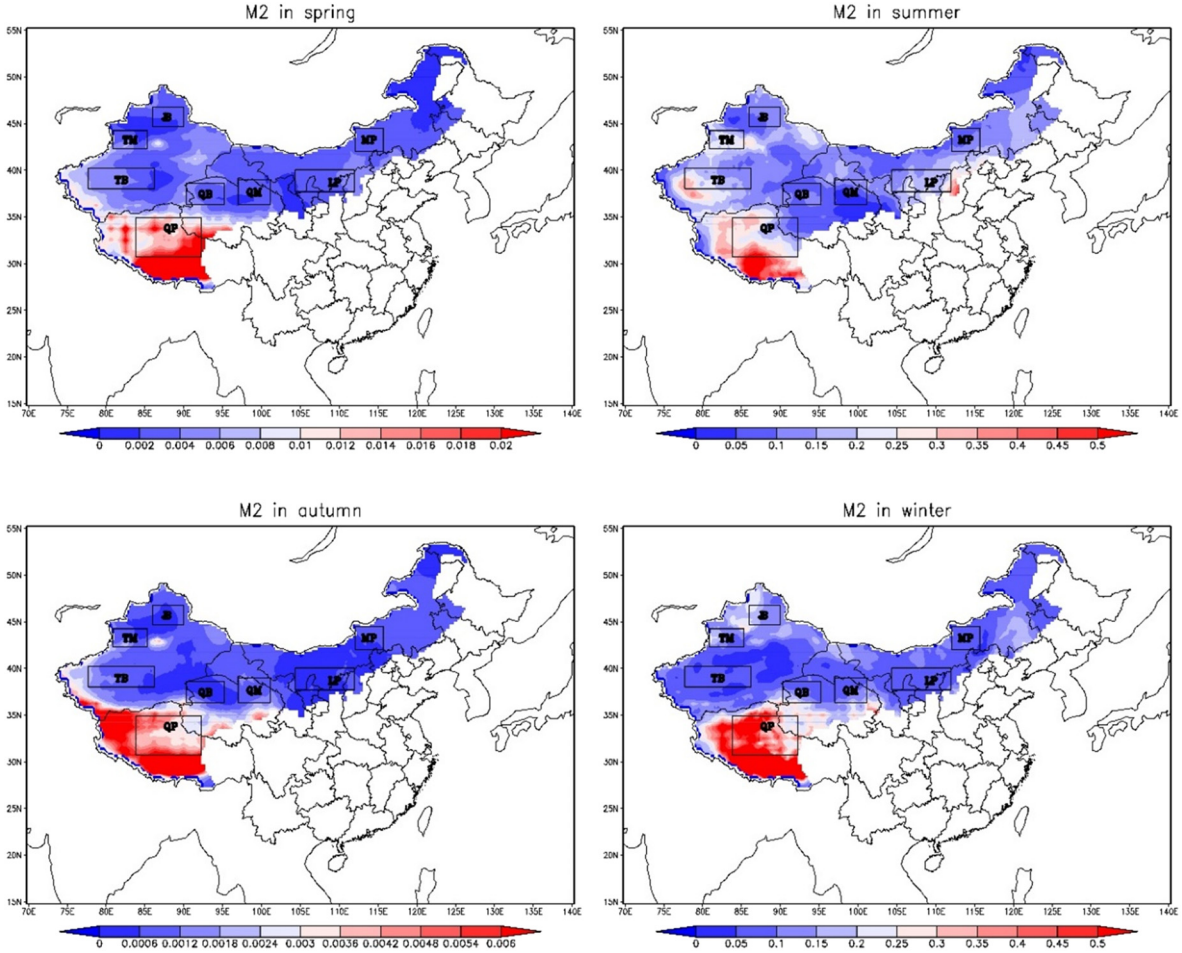


Fig. 5. Interannual variability statistic (M2) of seasonal temperature.

during winter than during summer. Topographically, the simulating skill in the basins is higher than that in the mountains and on the plateaus.

The models reproduce better precipitation and temperature during winter than during summer, consistent with previous studies reporting that models reproduce better precipitation during winter than during summer in East Asia (Ishizaki et al., 2012; Ma et al., 2015). IPCC AR5 (Kirtman et al., 2013) also indicates that the predictability of precipitation in China shows large seasonal differences, with reduced predictability during summer. The study area is mainly controlled by continental westerlies and high-pressure systems during winter. During summer, however, cyclones and subtropical high systems from tropical waters replace westerlies resulting in air flow conditions that are more complex than those during winter. In addition, convective and microphysical parameterization schemes also induce error sources in summer precipitation simulations (Chen and Frauenfeld, 2014). This is consistent with Zhao et al. (2008) and Madden and Shea (1978), who reported that climatic noise varies with pronounced seasonal variations; the largest variations typically occur during summer, and better predictability appears when limited climatic noise variance exists. Summarily, the seasonal prediction difference may be attributed to the influences of the climatic noise variations and atmospheric circulation. In contrast, some studies reported that the bias of the GCMs during winter was larger than that during summer, e.g., the Arctic by Chapman and Walsh (2007) and the Tibetan Plateau by Su et al. (2013). Su et al. (2013) used gauge data and commonly used metrics, e.g., correlation coefficient, RMSE, and bias in evaluating GCMs. In this study, an interpolated precipitation and temperature dataset and three skill scores were used. Different observational and statistical measures may lead to different conclusions.

The results also demonstrate that reproducibility significantly changes with topography. The precipitation and temperature reproducibility in basins is higher than in that of mountains and plateaus. This may be because the topography in mountains and on plateaus is more complex than that in basins. Previous studies found that the coarse resolution of climatic models makes it difficult to properly reproduce the atmospheric processes in spatially heterogeneous and complex terrain, because the complex topography can induce a local climate, e.g., *meso*scale mountain/valley winds (Chen et al., 2010; Su et al., 2013). In basins, precipitation is mainly produced by large-scale precipitation-bearing systems instead of convectional precipitation induced by complex orography. In addition, the physical scheme of GCM results shows better reproducibility for basins. The cloud microphysical schemes of GCMs are more applicable for weather systems on macro and meso scales (Bader et al., 2008). A weather system in a basin is relatively simple, with less local convectional weather systems at a small scale, comparing to mountains and plateaus. Su et al. (2013) indicated that systematic bias and significant problems over mountainous regions remain in the CMIP5 models, and further improvement in GCM performance is needed in regions with complex terrain.

4.2. Relationship between reproducibility and resolution

We found that the models with good precipitation simulating skill are not always suitable for temperature simulation. The only exception is that GISS-E2-H-CC and GISS-E2-R-CC are suitable for both

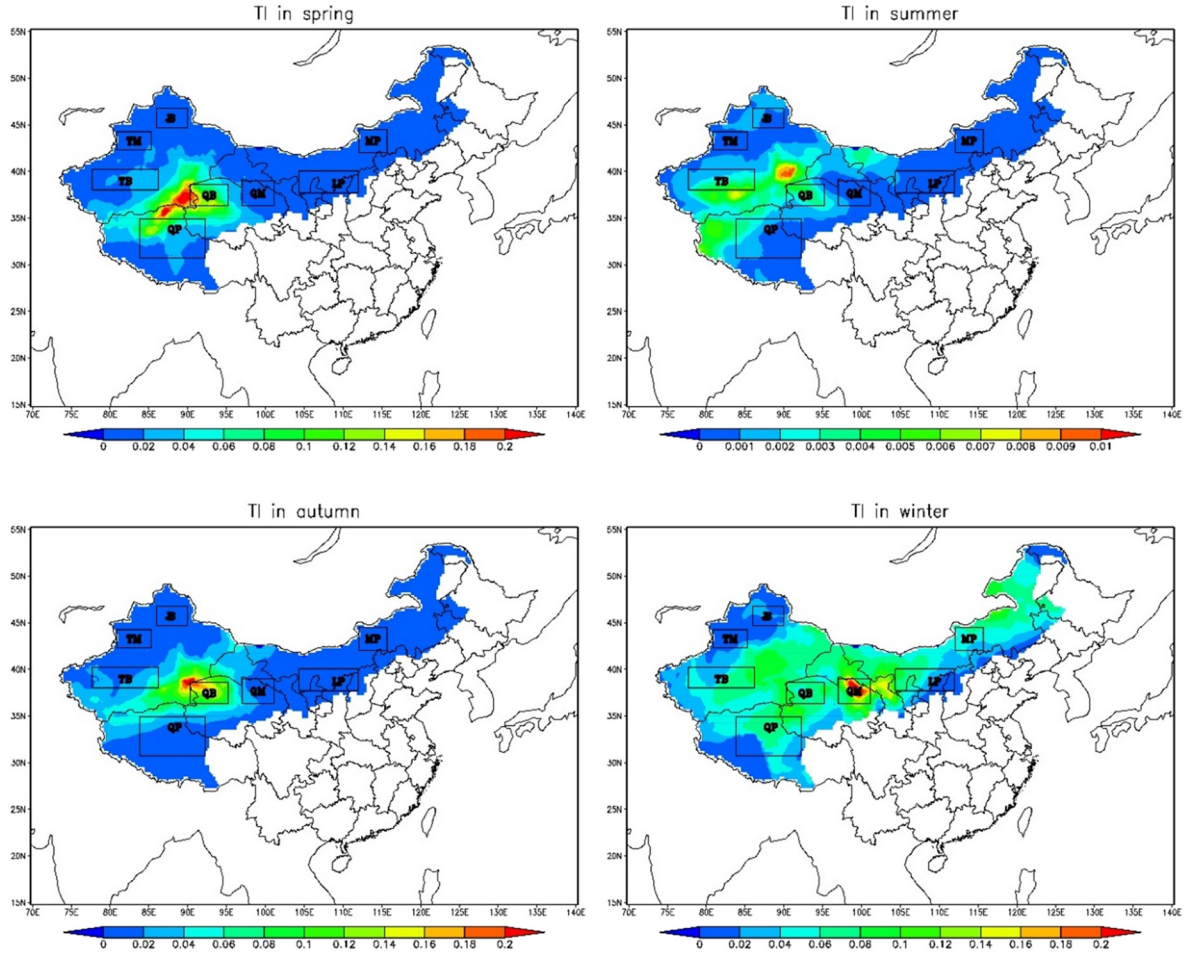


Fig. 6. Taylor skill score (TI) of seasonal precipitation.

precipitation and temperature. MIROC-ESM is the optimal model in basins for both precipitation and temperature. This result is consistent with the fact that no model fell in the best five for both temperature and precipitation, and likewise no model fell in the worst five for both, as reported by Mote and Salathe (2010). Among the top five models recommended by Zhao et al. (2014), only CSIRO-Mk3.6.0 for temperature produced results consistent with those in the present study. The conclusions of Zhao et al. (2014) were based on 28 models over China, while this study was based on 34 models over the arid and semiarid areas of China. The differences in the models and scale of the study area may have resulted in this inconsistency. It has been reported that models with relatively higher resolutions show better performance in China (W. Chen et al., 2011; Gao et al., 2006). However, the performance of the top models has no significant relationship with their resolution in this study, consistent with the conclusions of McMahon et al. (2015), Masson and Knutti (2011), Song and Zhou (2014), and Su et al. (2013).

4.3. Downscaling methods

A number of papers have previously indicated that GCMs are limited in their course resolution and systematic bias (Christensen et al., 2008; Park et al., 2016). The course resolution prevents the models from sufficiently representing regional climatic processes. The biases potentially grow when used for climate change simulations under global warming conditions; bias correction should be applied for each individual model (Christensen et al., 2008). Downscaling techniques are based on improving spatial resolution and reducing bias (Maurer and Hidalgo, 2008). The dynamical and statistical methods are two fundamental approaches. Dynamically downscaling nest high-resolution regional climate models (RCMs) in GCMs, it has definite physical meaning but is computationally expensive. Statistical downscaling is more straightforward and based on empirical relationships between the regional climate (predictands) and large-scale circulation (predictors) (Wilby et al., 1998). These methods range from simple scaling techniques to rather sophisticated approaches adopting probability mapping or weather generators. The principle is equalization of statistical characteristics (e.g. mean, variance, standard deviation, and residuals) between the GCMs and observations (Engen-Skaugen, 2007; Fang et al., 2015; Ho et al., 2012; Li et al., 2012). They were originally developed to downscale GCMs, but can also be used in bias correction (Teutschbein and Seibert, 2012), e.g. delta change, linear scaling (LS), local intensity scaling (LOCI), power transformation, variance scaling, and distribution mapping.

The delta change method constructs corresponding climatic scenarios, by combining the observations and mean climate of GCMs; for the detailed correcting process please see Section 2.2. The limitations are that, first, the bias is constant through time. Second, the mean, maximum, and minimum of the climatic variables are scaled, but the variance and spatial pattern of the climate is assumed constant in the future. Third, the temporal sequence of wet or dry days is constant (Diaz-Nieto and Wilby, 2005). However, it is the simplest statistical downscaling technique and can be rapidly applied to several GCMs. The output changes could be well represented by the downscaled results (Fowler et al., 2010).

The principle of a linear scaling approach is based on the differences between the observed and current simulated values. By definition, the mean of the corrected model outputs agrees with the mean of the observation. The precipitation and temperature are corrected by the ratio and

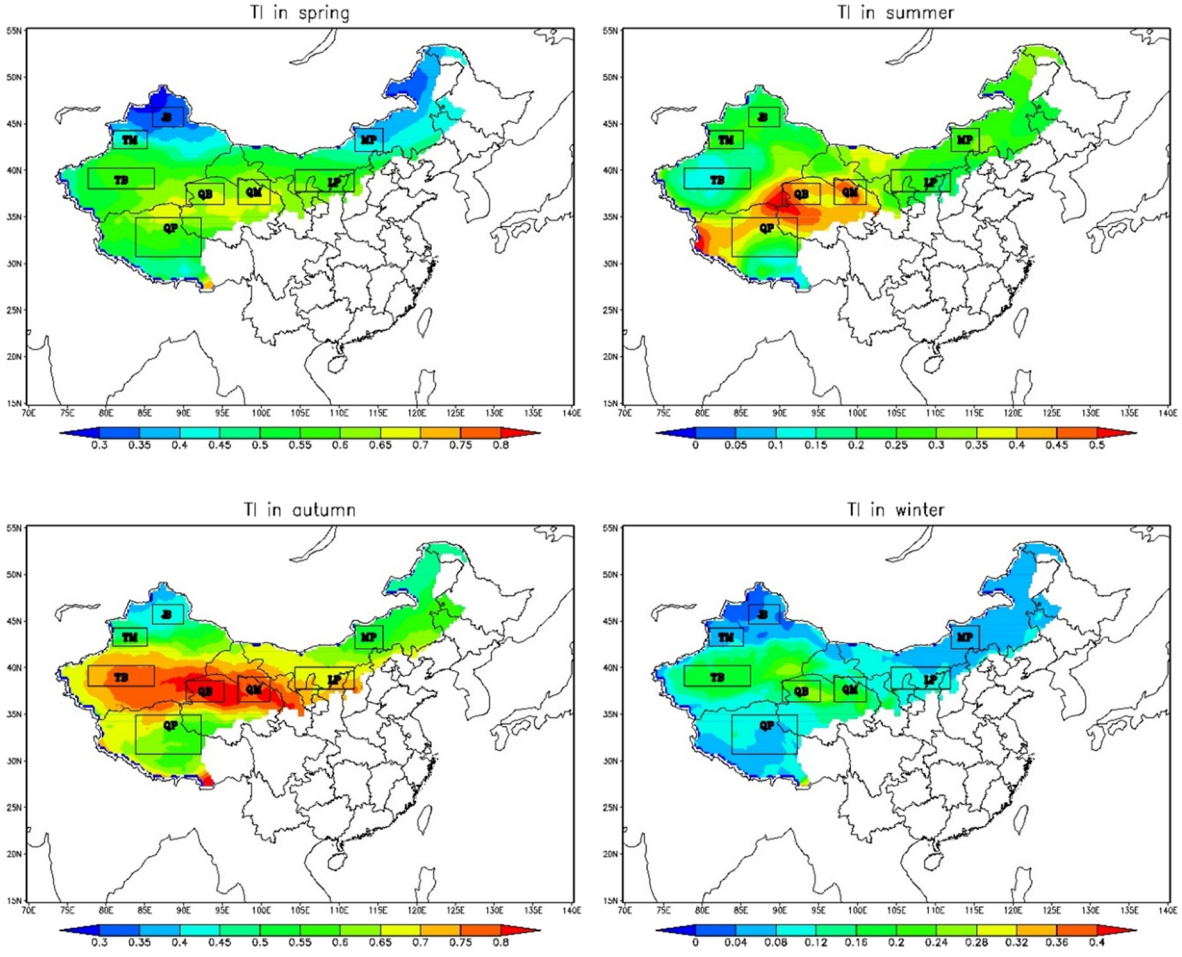


Fig. 7. Taylor skill score (TI) of seasonal temperature.

difference between long-term monthly mean observations and the current model simulation, respectively (Lenderink et al., 2007). The limitation is that it is not able to correct the biases in the wet-day frequency and intensity (Lenderink et al., 2007; Teutschbein and Seibert, 2012). On the basis of a linear scaling approach, Schmidli et al. (2006) went further and presented a local intensity scaling method, which adjusts the mean as well as the wet-day frequency and intensity, but only applies to a precipitation correction. The number of precipitation events in the models is first corrected by applying the calibrated model precipitation threshold. The precipitation is then corrected using a linear scaling factor, which is estimated based on the mean wet-day intensities. Leander and Buishand (2007) proposed a non-linear correction in an exponential form (a^*p^b) to adjust the variance in the precipitation, which cannot be corrected using the linear method. The root-finding algorithm is adopted to match the coefficient of variation in the corrected daily precipitation to that of the observed daily precipitation for each month, such that parameter b is determined. The aforementioned standard linear scaling parameter is then used to match the observed monthly mean precipitation with the intermediary series. The power function is only applicable for precipitation; the corresponding approach, termed variance scaling, to correct temperature was proposed by J. Chen et al. (2011). The means of the simulated temperature are first adjusted by linear scaling, and then the standard deviations are scaled. The standard deviation corrected temperature is finally shifted back using the corrected mean.

Quantile mapping (QM), a sophisticated approach developed by Panofsky and Brier (1968) and based on a transfer function derived from cumulative distribution functions (CDFs), is among the bias correction approaches that has been widely used in many recent studies (Chen et al., 2015; Ines and Hansen, 2006; Li et al., 2010; Piani et al., 2010). The method first finds the percentile values of a model future projection in a CDF for the training period, and then locates the observed values for the same CDF values. QM assumes that the future climatic distribution will remain similar to that in the reference period and does not change with time; only the mean changes. However, this assumption may not hold because of climatic nonstationarity (Milly et al., 2008; Mitchell and Jones, 2005). Li et al. (2010) improved traditional QM, and proposed a new QM termed EDCDFm. The most important improvement is that the method considered changes in future climatic distribution. Compared to other statistical downscaling methods (regression type models and weather classification), which are developed for small regions or specific sites, the new method is useful for a large spatial domain. In addition, the new method is more efficient in reducing biases than traditional QM for changing climates, particularly for the tails of the distribution. However, EDCDFm cannot guarantee that the corrected variable is assigned a positive value, which is illogical for some meteorological variables such as precipitation (Li et al., 2010). Wang and Chen (2014) proposed a method termed equiratio CDFm to avoid negative values for precipitation by using a multiplicative factor. However, the corrected precipitation will reach an unreasonably high value if the term of the inverse CDF is too small. Combining the previous two works, Miao et al. (2016) proposed a modified approach termed CNCDFm, which separately treated temperature and precipitation. Both parametric and nonparametric procedures were applied to estimate the CDF of the variables. This new method is effective in reducing the bias and correcting the entire distribution. However, the correcting process may violate the principle of the conservation of energy, because the priority is to increase the agreement between the GCM and observations, which usually neglects the underlying physical mechanisms, as in other statistical bias-correction methods (Ehret et al., 2012).

The studies related to method comparison showed that all bias correction approaches could reduce the bias to some extent, but the quality was strongly dependent on the method choice (Teutschbein and Seibert, 2012). Chen et al. (2013) showed that LOCI was better than LS; the distribution-based methods are consistently better than the mean-based methods over North America. Themeßl et al. (2011) compared seven approaches in the Alps region; the results showed that QM and LOCI had significant advantages compared to the multiple linear regression technique. QM showed the best performance.

Table 3
SS, M2 and TI of seasonal precipitation and temperature for each sub-region.

Indices	Regions	Precipitation						Temperature					
		Spring	Summer	Autumn	Winter	Seasonal mean	Mean	Spring	Summer	Autumn	Winter	Seasonal mean	Mean
SS	Study area	−0.52	−1.07	−1.03	**−0.43**	−0.76	–	−0.6	−0.58	−0.465	**−0.44**	−0.52	–
	JB	−0.74	−1.05	−1.00	−0.34	−0.78		**−0.48**	−0.56	**−0.29**	−0.35	**−0.42**	
	TB	**−0.19**	−0.94	**−0.55**	**−0.17**	**−0.46**	**−0.58**	−0.83	−0.59	−0.66	−0.43	−0.63	−0.52
	QB	−0.22	−0.82	−0.74	−0.25	−0.51		−0.61	**−0.45**	−0.48	−0.50	−0.51	
	TM	−0.99	−2.26	−2.03	−0.46	−1.44	−1.03	−0.57	−0.65	−0.40	**−0.33**	**−0.49**	**−0.49**
	QM	−0.48	−0.95	−0.65	−0.43	**−0.63**		−0.57	−0.46	−0.52	−0.43	**−0.49**	
	TP	−0.33	−1.19	−1.15	−0.50	−0.79		−0.77	−0.69	−0.63	−0.67	−0.69	
	MP	−0.52	**−0.68**	−0.85	−0.35	**−0.60**	−0.76	**−0.48**	−0.55	−0.34	−0.38	**−0.44**	−0.54
	LP	−0.71	−0.70	−1.24	−0.95	−0.90		−0.52	−0.65	−0.42	−0.41	−0.50	
M2	Study area	**0.72**	2.83	1.99	2.09	1.91	–	0.005	0.149	**0.001**	0.140	0.074	–
	JB	1.07	2.39	2.00	**0.70**	**1.54**		**0.002**	0.109	0.0006	0.150	0.065	
	TB	1.61	3.78	2.59	4.63	3.15	2.14	0.004	0.189	0.0010	**0.059**	0.063	**0.060**
	QB	0.23	1.85	1.61	3.16	1.71		0.003	0.092	0.0006	0.110	**0.051**	
	TM	0.95	10.41	4.92	0.93	4.31	2.48	0.003	0.206	0.0009	0.130	0.085	0.064
	QM	**0.18**	1.26	**0.31**	0.89	**0.66**		0.003	**0.073**	0.0010	0.096	**0.043**	
	TP	0.27	1.91	1.74	2.56	1.62		0.016	0.244	0.0041	0.438	0.176	
	MP	0.25	**0.49**	0.78	0.98	**0.62**	**1.29**	0.003	0.101	0.0006	0.072	**0.044**	0.094
	LP	1.16	0.57	1.97	2.85	1.64		0.003	0.176	**0.0005**	0.064	0.061	
TI	Study area	0.024	0.0013	0.024	**0.053**	0.030	–	**3.000**	0.295	0.645	0.111	1.010	–
	JB	0.004	0.0009	0.002	0.019	0.007		0.319	0.233	0.421	0.040	0.253	
	TB	0.021	0.0031	0.026	0.055	0.026	**0.036**	0.574	0.161	0.761	0.182	0.419	0.400
	QB	**0.102**	**0.0031**	**0.127**	0.069	**0.075**		0.642	**0.434**	**0.820**	**0.213**	**0.527**	
	TM	0.005	0.0005	0.004	0.010	0.005	0.022	0.435	0.254	0.522	0.058	0.317	**0.412**
	QM	0.013	0.0006	0.016	**0.130**	**0.040**		**0.646**	0.419	0.793	0.168	**0.506**	
	TP	0.039	0.0011	0.013	0.046	**0.025**		0.571	0.326	0.619	0.084	0.400	
	MP	0.008	0.0005	0.005	0.051	0.016	0.018	0.399	0.278	0.533	0.052	0.315	0.373
	LP	0.002	0.0004	0.002	0.043	0.012		0.577	0.257	0.694	0.092	**0.405**	

Note: the bold types were the best values. The first line for each index was compared between seasons for the entire study area. The column "Spring", "Summer", "Autumn", "Winter", and "Seasonal mean" were compared between sub-regions for each season, or seasonal mean, and the values for the study area was not included. The column "Mean" was compared among three kinds of topography: the basin, mountain, and plateau.

Teutschbein and Seibert (2012) concluded that quantile mapping performed best in five meso-scale catchments in Sweden. Gudmundsson et al. (2012) indicated that nonparametric transformations outperformed distribution-derived transformation and parametric transformation in Norway. Lafon et al. (2013) obtained similar results, that gamma-based quantile mapping performed better than linear, nonlinear, and empirical distribution-based quantile mapping over north America. Van Roosmalen et al. (2011) analyzed the mean hydrological statistics and shows that the delta change and distribution mapping were not significantly different in the projection of changes in mean hydrological responses for a Denmark catchment.

This brief review showed that each approach has strengths and weaknesses (Fowler et al., 2010). It is difficult to assess the relative performance of different downscaling methods; there is no consistently superior downscaling method. Li et al. (2010) indicated that there is no panacea, because the statistical downscaling approach is generally developed for a specific purpose or location. The main objective of this study was to compare the model simulation skills among different topographies and seasons, and select the applicable models for the study area. Using the same method to downscale all models could reduce the uncertainties brought by different methods; thus, the results might be more comparable. However, only mean precipitation and temperature were analyzed, while other indices such as variance and wet/dry day were not discussed in this study. The delta method is the simplest technique suitable to downscale the mean value of climatic variables. This method ensures that the downscaled results of the amount and tendency of climatic variables are quite close to those of the observations, which is sufficient to meet our need.

4.4. Factors influencing climate change

Climate is a complex physical system; a change is influenced by physical, chemical, and biological processes.

The Fourth Assessment Report of IPCC concluded that most of the observed increase in global mean temperatures is very likely a resulted of the observed increase in anthropogenic greenhouse gas (GHG) concentrations since the mid-twentieth century (IPCC, 2007). Human activities (e.g., fossil fuel burning, agricultural activities such as growing rice,

Table 4
Weights and ranks in each season for regions according to models' performance.

Region	Season	Precipitation					Temperature				
		Rank of SS	Rank of M2	Rank of TI	sum of the ranks	Weights	Rank of SS	Rank of M2	Rank of TI	sum of the ranks	Weights
JB	Spring	2	2	2	6	0.240	3	2	2	7	0.211
	Summer	4	4	4	12	0.120	4	3	3	10	0.148
	Autumn	3	3	3	9	0.160	1	1	1	3	**0.493**
	Winter	1	1	1	3	**0.480**	2	4	4	10	0.148
TM	Spring	2	2	2	6	0.240	3	2	2	7	0.235
	Summer	4	4	4	12	0.120	4	4	3	11	0.149
	Autumn	3	3	3	9	0.160	2	1	1	4	**0.411**
	Winter	1	1	1	3	**0.480**	1	3	4	8	0.205
TB	Spring	2	1	3	6	**0.294**	4	2	2	8	0.220
	Summer	4	3	4	11	0.160	2	4	4	10	0.176
	Autumn	3	2	2	7	0.252	3	1	1	5	**0.352**
	Winter	1	4	1	6	**0.294**	1	3	3	7	0.252
QB	Spring	1	1	2	4	**0.404**	4	2	2	8	0.205
	Summer	4	3	4	11	0.147	1	3	3	7	0.235
	Autumn	3	2	1	6	0.269	2	1	1	4	**0.411**
	Winter	2	4	3	9	0.180	3	4	4	11	0.149
QM	Spring	2	1	3	6	0.281	4	2	2	8	0.223
	Summer	4	4	4	12	0.141	2	3	3	8	0.223
	Autumn	3	2	2	7	0.241	3	1	1	5	**0.356**
	Winter	1	3	1	5	**0.337**	1	4	4	9	0.198
QP	Spring	1	1	2	4	**0.411**	4	2	2	8	0.187
	Summer	4	3	4	11	0.149	3	3	3	9	0.166
	Autumn	3	2	3	8	0.205	1	1	1	3	**0.498**
	Winter	2	4	1	7	0.235	2	4	4	10	0.149
MP	Spring	2	1	2	5	**0.346**	3	2	2	7	0.211
	Summer	3	2	4	9	0.192	4	4	3	11	0.134
	Autumn	4	3	3	10	0.173	1	1	1	3	**0.491**
	Winter	1	4	1	6	0.288	2	3	4	9	0.164
LP	Spring	2	2	3	7	0.262	3	2	2	7	0.235
	Summer	1	1	4	6	**0.305**	4	4	3	11	0.149
	Autumn	4	3	2	9	0.204	2	1	1	4	**0.411**
	Winter	3	4	1	8	0.229	1	3	4	8	0.205
Basin	Spring	2	1	3	6	**0.294**	4	2	2	8	0.208
	Summer	4	3	4	11	0.160	3	4	3	10	0.167
	Autumn	3	2	1	6	**0.294**	2	1	1	4	**0.417**
	Winter	1	4	2	7	0.252	1	3	4	8	0.208
Plateau	Spring	2	1	3	6	0.267	3	2	2	7	0.211
	Summer	4	4	4	12	0.133	4	3	3	10	0.148
	Autumn	3	3	2	8	0.200	1	1	1	3	**0.493**
	Winter	1	2	1	4	**0.400**	2	4	4	10	0.148
Mountain	Spring	1	1	2	4	**0.414**	4	2	2	8	0.208
	Summer	3	2	4	9	0.184	3	4	3	10	0.167
	Autumn	4	3	3	10	0.166	2	1	1	4	**0.417**
	Winter	2	4	1	7	0.237	1	3	4	8	0.208
Study area	Spring	2	1	2	5	**0.329**	4	2	1	7	**0.252**
	Summer	4	4	4	12	0.137	3	4	3	10	0.176
	Autumn	3	2	3	8	0.205	2	1	2	5	**0.352**
	Winter	1	3	1	5	**0.329**	1	3	4	8	0.220

The bold is the biggest weight in each region.

Table 5
Weights and ranks of seasonal mean for regions according to models' performance.

Region	Precipitation					Temperature				
	Rank of SS	Rank of M2	Rank of TI	Sum of the ranks	Weights	Rank of SS	Rank of M2	Rank of TI	sum of the ranks	Weights
JB	5	3	7	15	0.099	1	6	8	15	0.102
TM	8	8	8	24	0.062	3	7	6	16	0.096
TB	1	7	3	11	**0.135**	7	5	3	15	0.102
QB	2	6	1	9	**0.165**	6	3	1	10	**0.153**
QM	4	2	2	8	**0.186**	4	1	2	7	**0.218**
QP	6	4	4	14	0.106	8	8	5	21	0.073
MP	3	1	5	9	**0.165**	2	2	7	11	**0.139**
LP	7	5	6	18	0.082	5	4	4	13	0.118
Basin	1	2	1	4	**0.462**	2	1	2	5	0.356
Plateau	2	1	3	6	0.308	3	3	3	9	0.198
Mountain	3	3	2	8	0.231	1	2	1	4	**0.446**

The bolds in the first group indicated the weight of top three. The bold in the second group indicated the maximun weight.

nitrogen fertilizer application, and rearing of livestock) have increased the amount of GHGs (carbon dioxide, methane, and nitrous oxide) in the atmosphere. Emissions of sulfate aerosols and smaller forcing including stratospheric ozone depletion, black and organic carbon aerosols, and jet contrails also affect the climate. In addition, the movement of four important elements for life (carbon, oxygen, nitrogen, and phosphorus) and major elements with toxic properties (e.g., lead, mercury, arsenic, and cadmium) have been augmented by human activities such as mining and industrial production. These changes in the Earth's chemistry have resulted in global climate change (Hunter, 2017).

It is known that vegetation influences climate and that phenology is an important factor that regulates vegetation feedback to the atmosphere and climate. The influencing mechanisms include albedo, surface roughness length, canopy conductance, photosynthetic and CO_2 fluxes, and biogenic volatile organic compound (BVOC) fluxes. Albedo plays an important role in the surface energy budget (Bonan, 2008) and changes with the development of the canopy and ecosystem types (Richardson et al., 2013). Surface roughness length regulates land surface energy fluxes by influencing the coupling degree between the land surface and the atmosphere (Pitman, 2010). Canopy conductance mediates CO_2 uptake rates and transpiration (Richardson et al., 2013). Phenology affects variations in surface–atmosphere CO_2 fluxes through photosynthesis and spatial and interannual patterns of C uptake, which thus influence the global atmosphere (Richardson et al., 2010). BVOCs control a variety of feedback mechanisms within the climatic system.

Table 6
Weights of the models for regions according to their precipitation reproducibility.

No.	Model	Region											
		JB	TM	TB	QB	QM	QP	MP	LP	Basin	Plateau	Mountain	Study area
1	ACCESS1-0	0.021	0.032	0.030	**0.043**	0.015	**0.044**	0.020	0.015	0.039	0.020	0.020	0.028
2	ACCESS1-3	0.023	0.028	0.020	0.022	0.014	**0.044**	0.040	**0.063**	0.018	**0.082**	0.018	0.025
3	bcc-csm1-1	0.023	0.016	0.030	0.018	0.025	0.024	0.015	0.027	0.022	0.020	0.020	0.021
4	bcc-csm1-1-m	0.013	0.014	0.023	0.020	0.022	0.033	0.018	0.013	0.019	0.016	0.018	0.015
5	BNU-ESM	0.015	0.012	0.020	0.017	0.030	0.015	0.024	0.032	0.016	0.019	0.016	0.014
6	CanESM2	0.015	0.011	0.018	0.013	0.014	0.012	0.014	0.010	0.013	0.012	0.012	0.012
7	CCSM4	0.014	0.022	0.027	0.025	**0.044**	0.042	0.022	0.022	0.023	0.030	0.039	0.034
8	CESM1-BGC	0.024	0.022	0.025	0.028	0.031	**0.127**	0.024	0.030	0.027	**0.052**	0.023	0.035
9	CESM1-CAM5	0.015	0.015	0.033	0.017	0.016	0.019	**0.048**	**0.042**	0.022	0.038	0.017	0.022
10	CESM1-FASTCHEM	0.016	0.016	0.019	0.031	0.022	**0.082**	0.021	0.023	0.023	0.032	0.021	0.028
11	CESM1-WACCM	0.013	0.011	0.035	0.021	0.013	0.016	**0.050**	**0.053**	0.021	0.031	0.013	0.016
12	CSIRO-Mk3-6-0	0.018	0.014	0.028	0.016	0.013	0.012	0.025	0.013	0.020	0.013	0.015	0.013
13	EC-EARTH	0.011	0.015	0.015	0.016	0.018	**0.060**	0.016	0.013	0.013	0.017	0.016	0.013
14	FGOALS-g2	**0.049**	0.038	**0.042**	0.032	**0.047**	0.022	**0.046**	**0.044**	0.037	0.039	**0.056**	**0.051**
15	FIO-ESM	**0.041**	0.016	0.032	0.026	0.039	0.022	0.012	0.012	0.031	0.014	0.023	0.020
16	GFDL-ESM2G	0.011	0.011	0.016	0.016	0.016	0.014	0.019	0.012	0.014	0.013	0.014	0.012
17	GFDL-ESM2M	0.011	0.010	0.015	0.013	0.014	0.012	0.016	0.011	0.013	0.012	0.013	0.012
18	GISS-E2-H-CC	0.021	0.041	**0.041**	0.031	**0.065**	0.021	**0.052**	**0.036**	0.035	0.041	**0.059**	**0.051**
19	GISS-E2-H	0.034	**0.054**	**0.042**	0.039	**0.044**	0.021	**0.046**	0.022	**0.044**	0.028	0.045	0.047
20	GISS-E2-R-CC	0.034	**0.079**	**0.042**	0.040	**0.077**	0.020	**0.066**	0.027	0.043	0.033	**0.084**	**0.049**
21	GISS-E2-R	0.028	**0.102**	0.029	0.034	**0.044**	0.017	0.037	0.027	0.038	0.023	**0.069**	0.043
22	HadGEM2-AO	0.029	0.032	0.037	0.034	0.026	0.025	0.017	0.025	0.030	0.022	0.028	0.027
23	HadGEM2-CC	0.022	0.035	0.038	0.030	0.028	0.027	0.012	0.013	0.025	0.015	0.035	0.021
24	HadGEM2-ES	0.016	0.033	0.038	0.025	0.035	0.028	0.016	0.023	0.023	0.021	0.045	0.023
25	inmcm4	**0.041**	0.045	0.031	0.022	0.021	0.025	0.041	0.015	0.031	0.021	0.034	0.029
26	IPSL-CM5A-LR	0.016	0.016	0.035	0.030	0.029	0.017	0.019	0.012	0.032	0.016	0.023	0.019
27	IPSL-CM5A-MR	0.022	0.029	0.026	0.026	0.036	0.015	0.025	0.012	0.026	0.015	0.029	0.020
28	IPSL-CM5B-LR	0.012	0.012	0.019	0.015	0.021	0.013	0.019	0.010	0.015	0.012	0.017	0.013
29	MIROC4h	0.040	**0.102**	**0.042**	**0.047**	0.040	0.019	0.037	**0.202**	**0.051**	**0.057**	0.047	**0.078**
30	MIROC5	0.038	0.021	0.039	**0.049**	0.026	0.020	0.032	0.014	**0.058**	0.018	0.021	0.034
31	MIROC-ESM-CHEM	**0.077**	0.022	0.034	**0.064**	0.030	0.034	**0.046**	**0.036**	**0.049**	**0.064**	0.027	0.042
32	MIROC-ESM	**0.119**	**0.047**	0.033	**0.075**	**0.049**	0.042	**0.050**	0.035	**0.064**	**0.088**	**0.049**	**0.084**
33	NorESM1-ME	**0.090**	0.013	0.023	0.028	0.018	0.026	0.026	0.031	0.030	0.034	0.018	0.023
34	NorESM1-M	0.028	0.013	0.023	0.038	0.018	0.031	0.028	0.024	0.033	0.032	0.017	0.028

Top five weights for each region were bolded, indicating the top five models.

The products of BVOC reactions influence radiative forcing, which in turn increases the concentration of cloud condensation nuclei through aerosol production, reduces the oxidation capacity of the troposphere, and changes ozone levels. The aforementioned processes influence the surface energy balance and the composition and structure of the surface boundary layer, and in turn the climate (Curci et al., 2009; Laothawornkitkul et al., 2010; Mentel et al., 2009). Land cover change affects the climate via the mechanism of surface albedo variation and emissions of carbon dioxide (Bolin et al., 2000; Matthews et al., 2004).

Natural climatic forcing such as intermittent volcanic eruptions and solar variability have received considerable attention. The global surface temperature response to volcanic eruptions is cooling, resulting from increased absorption and reflection of incoming shortwave radiation by stratospheric aerosols (Shindell et al., 2003). The most obvious impact of the sun is its influence on the Earth's radiation budget through variations in total solar irradiance (TSI). Variation in solar ultraviolet radiation affect stratospheric ozone, leading to associated temperature variations. Solar forcing contributed to an increase in global temperatures during the early part of the twentieth century, but played only a minor role during the latter part (Gray et al., 2010). It also has been indicated in previous studies that solar activity may result in climate change (Gray et al., 2010; Kovalenko and Zherebtsov, 2014; Mufti and Shah, 2011). The temperature increases during the 12th and 13th centuries and decreases during the Little Ice Age corresponded with high and low solar activity, respectively. After termination of the Maunder minimum, a general increase in the solar activity level began and from 1900 to 2000 the world climate has warmed. Parameters of the solar wind and interplanetary magnetic field that determine geomagnetic activity and solar space rays are key in assessing the solar activity influence on the climatic characteristics of the troposphere (Kovalenko and Zherebtsov, 2014). Periodic variations in the Earth's orbit and tilt also affect the distance from the Earth to the sun and the distribution of sunlight over the Earth's surface, which in turn influences the climate (Ganguly, 2011). In addition, oceans are the Earth's largest thermal reservoirs (Abraham et al., 2013). The influence of solar activity on climate is conditioned by the circulation process in the surface layer of the ocean and atmosphere and the effectiveness of heat exchange in ocean-atmosphere-land systems affected the climate to a large extent. The decline in sea ice cover also enhances the impact of global warming, as it results in an albedo decline, and air temperatures in turn increase (Zherebtsov et al., 2015). With population growth, industrialization, and urbanization, new processes (e.g., dynamic vegetation, biogeochemistry, and aerosol chemistry) will result in more complex climate change. Related research will be more interesting and challenging.

4.5. Limitations of this study

Climatic variables such as precipitation, temperature, humidity, radiation, wind speed, etc. may affect the eco-hydrological process. Among which, changes in precipitation and temperature have the dominant influence, because regional hydrology will change in turn to a large extent (Rong et al., 2013; Zhao and Yu, 2013). Precipitation controls the water input of the eco-hydrological system and has a direct effect on vegetation, soil moisture, and runoff. Temperature has a significant influence on evapotranspiration and the moisture cycle. Precipitation and temperature are the most important inputs for a hydrological model. It has been indicated that accurate inputs are the foundation for reliable hydrologic modeling (Duncan et al., 1993), and are often more critical than the choice of model (Gan et al., 1997). For this reason, only precipitation and temperature were evaluated here to facilitate further research related to impacts of climate change on the eco-hydrology process using hydrological models in arid and

Table 7
Weights of the models for regions according to their temperature reproducibility.

No.	Model	Region											
		JB	TM	TB	QB	QM	QP	MP	LP	Basin	Plateau	Mountain	Study area
1	ACCESS1-0	0.026	0.021	0.012	0.012	0.012	0.013	0.017	0.019	0.014	0.016	0.014	0.013
2	ACCESS1-3	0.025	0.022	0.015	**0.057**	0.030	**0.171**	0.018	0.021	0.023	0.028	0.023	0.026
3	bcc-csm1-1	0.023	0.024	0.024	0.022	0.022	**0.086**	0.047	0.022	0.028	**0.073**	0.028	**0.041**
4	bcc-csm1-1-m	0.019	0.013	0.011	0.015	0.015	0.034	0.013	0.012	0.013	0.016	0.014	0.013
5	BNU-ESM	0.014	0.012	0.013	0.023	0.026	0.026	0.013	0.012	0.013	0.013	0.013	0.013
6	CanESM2	0.031	0.016	0.023	0.022	0.016	0.009	0.028	0.013	0.030	0.011	0.015	0.015
7	CCSM4	0.016	0.018	0.012	0.024	**0.045**	0.012	0.020	0.018	0.016	0.014	0.027	0.018
8	CESM1-BGC	0.023	0.026	0.031	0.027	0.040	0.015	0.024	0.028	0.039	0.021	0.036	0.031
9	CESM1-CAM5	**0.040**	0.020	0.026	0.012	0.018	0.009	0.026	**0.058**	0.020	0.014	0.021	0.018
10	CESM1-FASTCHEM	0.023	0.018	0.015	0.020	0.021	0.011	0.020	0.013	0.019	0.013	0.021	0.017
11	CESM1-WACCM	0.035	0.036	0.019	0.014	0.017	0.011	**0.059**	0.038	0.022	0.018	0.029	0.020
12	CSIRO-Mk3-6-0	0.027	0.021	**0.044**	**0.068**	**0.061**	**0.122**	0.051	**0.043**	**0.045**	**0.190**	0.027	**0.083**
13	EC-EARTH	0.028	0.022	0.031	**0.077**	**0.093**	0.019	0.019	0.023	0.035	0.024	0.032	0.040
14	FGOALS-g2	0.017	0.016	0.037	0.038	0.021	0.048	0.018	0.014	0.029	0.026	0.017	0.026
15	FIO-ESM	0.031	0.021	0.020	0.019	0.018	0.021	0.031	**0.123**	0.025	**0.050**	0.023	0.031
16	GFDL-ESM2G	0.031	0.022	0.014	0.022	0.029	**0.053**	0.013	0.014	0.020	0.017	0.022	0.019
17	GFDL-ESM2M	0.038	0.031	0.012	0.015	0.035	0.032	0.026	0.018	0.017	0.031	0.031	0.022
18	GISS-E2-H-CC	**0.048**	**0.059**	0.033	**0.060**	0.043	**0.061**	**0.059**	0.031	**0.058**	**0.086**	**0.062**	**0.106**
19	GISS-E2-H	0.034	**0.040**	**0.046**	0.044	0.043	0.021	0.023	0.019	0.039	0.020	0.036	0.040
20	GISS-E2-R-CC	**0.041**	**0.040**	0.039	0.041	0.038	0.023	**0.079**	0.033	**0.045**	0.043	**0.046**	**0.061**
21	GISS-E2-R	0.030	0.031	0.024	0.038	0.036	0.023	0.025	0.019	0.035	0.022	0.034	0.031
22	HadGEM2-AO	0.021	0.023	0.011	0.014	0.015	0.009	0.018	0.012	0.013	0.010	0.019	0.012
23	HadGEM2-CC	**0.041**	**0.095**	0.042	0.041	0.019	0.011	0.013	0.012	**0.064**	0.010	0.035	0.022
24	HadGEM2-ES	0.022	**0.040**	**0.066**	0.023	0.023	0.013	0.013	0.019	0.044	0.014	0.031	0.025
25	inmcm4	0.025	0.019	0.013	0.011	0.012	0.008	0.014	0.011	0.013	0.009	0.013	0.012
26	IPSL-CM5A-LR	0.022	0.027	0.014	0.013	0.014	0.011	**0.054**	0.015	0.016	0.015	0.019	0.017
27	IPSL-CM5A-MR	0.032	0.019	0.019	0.018	0.022	0.011	0.047	0.016	0.021	0.014	0.022	0.020
28	IPSL-CM5B-LR	0.028	0.015	0.012	0.012	0.013	0.020	0.020	0.012	0.015	0.015	0.014	0.015
29	MIROC4h	0.035	0.032	**0.082**	0.048	0.039	0.019	0.033	0.035	0.036	0.031	**0.038**	0.039
30	MIROC5	**0.073**	**0.065**	0.030	0.038	**0.045**	0.009	**0.056**	**0.070**	**0.051**	0.020	**0.073**	0.034
31	MIROC-ESM-CHEM	0.017	0.019	0.035	0.025	0.025	0.014	0.014	0.011	0.024	0.011	0.023	0.018
32	MIROC-ESM	0.034	**0.062**	**0.133**	**0.050**	**0.048**	0.011	0.026	0.016	**0.072**	0.015	**0.089**	0.037
33	NorESM1-ME	0.020	0.020	0.019	0.015	0.018	0.022	0.018	0.019	0.019	0.021	0.020	0.022
34	NorESM1-M	0.031	0.036	0.021	0.019	0.027	0.023	0.044	**0.163**	0.028	**0.068**	0.034	**0.043**

Top five weights for each region were bolded, indicating the top five models.

semi-arid areas of China. In future work, other variables such as radiation and humidity can be considered.

5. Conclusions

We investigated the precipitation and temperature simulating capabilities of 34 GCMs from 1961 to 1999 for eight topographically representative sub-regions in the arid and semiarid region of China. Model performances were greatly influenced by season and topography. Seasonally, the reproducibility during winter was better than that during summer and better during spring than that during autumn for precipitation. The weights were 0.329 (0.24–0.411) during spring, 0.137 (0.12–0.305) during summer, 0.205 (0.16–0.269) during autumn and 0.329 (0.18–0.48) during winter at the scale of the whole study area (or sub-region). The simulating skill during autumn was better than during spring for temperature; the weight during autumn was 0.1 and 0.132–0.311 higher than during spring for the whole study area and sub-regions, respectively. The temperature reproducibility during winter for the whole study area and four sub-regions (JB, TM, MP, and LP) was better than that during summer; the weights were 0.044 and 0.03–0.056 higher, respectively. Topographically, the precipitation reproducibility in the QM, QB, MP, and TB was better; the weights were 0.186, 0.165, 0.165, and 0.135, respectively. The temperature reproducibility in the QM, QB, and MP was better; the weights were 0.218, 0.153, and 0.139, respectively. On the whole, the precipitation reproducibility in the basins was the best, followed by plateaus and mountains; the weights were 0.462, 0.308, and 0.231, respectively. The temperature reproducibility in the mountains was the best, followed by basins and plateaus; the weights were 0.446, 0.356, and 0.198, respectively. The top five models were MIROC-ESM, MIROC4h, FGOALS-g2, GISS-E2-H-CC, GISS-E2-R-CC for precipitation, and GISS-E2-H-CC, CSIRO-Mk3-6-0, GISS-E2-R-CC, NorESM1-M, bcc-csm1-1 for temperature of the whole study area. For precipitation, the results recommended MIROC-ESM for basins and plateaus, and GISS-E2-R-CC for mountains. For temperature, MIROC-ESM for basins and mountains and CSIRO-Mk3-6-0 for plateaus were recommended. The top five models for each sub-region were also recommended. The reproducibility of precipitation and temperature, as well as the recommended GCMs, have reference value in future studies related to climate change and their impacts on eco-hydrological simulation.

Acknowledgments

This work was supported by The National Key Research and Development Program of China (2016YFC0501603), The National Natural Science Foundation of China (41501301), and the Chinese Academy of Sciences "Light of West China" Program. We would like to thank all the modeling groups for making available the CMIP5 multi-model datasets.

Author contributions

Yanfen Yang and Lei Bai analyzed the results; Yanfen Yang wrote the paper; and Lei Bai, Bing Wang, Jing Wu, and Suhua Fu provided suggestions for the analysis and revised the paper.

References

Abraham, J.P., Baringer, M., Bindoff, N.L., Boyer, T., Cheng, L.J., Church, J.A., et al., 2013. A review of global ocean temperature observations: implications for ocean heat content estimates and climate change. Rev. Geophys. 51, 450–483.

Anandhi, A., Frei, A., Pierson, D.C., Schneiderman, E.M., Zion, M.S., Lounsbury, D., et al., 2011. Examination of change factor methodologies for climate change impact assessment. Water Resour. Res. 47, 341–351.

Bader, D.C., Covey, C., Gutowski, W.J., Held, I.M., Kunkel, K.E., Miller, R.L., et al., 2008. Climate models: an assessment of strengths and limitations. A Report by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research. Department of Energy, Office of Biological and Environmental Research, Washington, D.C., USA 124 pp.

Bolin, B., Sukumar, R., Ciais, P., Cramer, W., Jarvis, P., Kheshgi, H., et al., 2000. Global perspective. In: Watson, R.T., et al. (Eds.), Land Use, Land-use Change, and Forestry: A Special Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK.

Bonan, G.B., 2008. Forests and climate change: forcings, feedbacks, and the climate benefits of forests. Science 320, 1444–1449.

Chapman, W.L., Walsh, J.E., 2007. Simulations of Arctic temperature and pressure by global coupled models. J. Clim. 20, 609–632.

Chen, X., 2010. Physical Geography of Arid Land in China. Science Press, Beijing (in Chinese).

Chen, L., Frauenfeld, O.W., 2014. A comprehensive evaluation of precipitation simulations over China based on CMIP5 multimodel ensemble projections. J. Geophys. Res. Atmos. 119, 5767–5786.

Chen, Y.N., Zilliacus, H., Li, W.H., Zhang, H.F., Chen, Y.P., 2006. Ground-water level affects plant species diversity along the lower reaches of the Tarim river, Western China. J. Arid Environ. 66, 231–246.

Chen, H.M., Zhou, T.J., Neale, R.B., Wu, X.Q., Zhang, G.J., 2010. Performance of the new NCAR CAM3.5 in East Asian summer monsoon simulations: sensitivity to modifications of the convection scheme. J. Clim. 23, 3657–3675.

Chen, J., Brissette, F.P., Leconte, R., 2011. Uncertainty of downscaling method in quantifying the impact of climate change on hydrology. J. Hydrol. 401, 190–202.

Chen, W., Jiang, Z., Li, L., 2011. Probabilistic projections of climate change over China under the SRES A1B scenario using 28 AOGCMs. J. Clim. 4741–4756.

Chen, Y.N., Ye, Z.X., Shen, Y.J., 2011c. Desiccation of the Tarim River, Xinjiang, China, and mitigation strategy. Quat. Int. 244, 264–271.

Chen, J., Brissette, F.P., Chaumont, D., Braun, M., 2013. Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America. Water Resour. Res. 49, 4187–4205.

Chen, J., Brissette, F.P., Lucas-Picher, P., 2015. Assessing the limits of bias-correcting climate model outputs for climate change impact studies. J. Geophys. Res. Atmos. 120, 1123–1136.

Christensen, J.H., Boberg, F., Christensen, O.B., Lucas-Picher, P., 2008. On the need for bias correction of regional climate change projections of temperature and precipitation. Geophys. Res. Lett. 35, 229–237.

Curci, G., Beekmann, M., Vautard, R., Smiatek, G., Steinbrecher, R., Theloke, J., et al., 2009. Modelling study of the impact of isoprene and terpene biogenic emissions on European ozone levels. Atmos. Environ. 43, 1444–1455.

Diaz-Nieto, J., Wilby, R.L., 2005. A comparison of statistical downscaling and climate change factor methods: impacts on low flows in the River Thames, United Kingdom. Clim. Chang. 69, 245–268.

Duan, Q.Y., Phillips, T.J., 2010. Bayesian estimation of local signal and noise in multimodel simulations of climate change. J. Geophys. Res. Atmos. 115 (-).

Duncan, M.R., Austin, B., Fabry, F., Austin, G.L., 1993. The effect of gauge sampling density on the accuracy of streamflow prediction for rural catchments. J. Hydrol. 142, 445–476.

Ehret, U., Zehe, E., Wulfmeyer, V., 2012. Should we apply bias correction to global and regional climate model data? Hydrol. Earth Syst. Sc 16, 3391–3404 (19,1(2015-01-21)).

Engen-Skaugen, T., 2007. Refinement of dynamically downscaled precipitation and temperature scenarios. Clim. Chang. 84, 365–382.

Fang, G.H., Yang, J., Chen, Y.N., Zammit, C., 2015. Comparing bias correction methods in downscaling meteorological variables for hydrologic impact study in an arid area in China. Hydrol. Earth Syst. Sci. 11, 2547–2559 (19,6(2015-06-02)).

Feng, Q., Liu, W., Si, J.H., Su, Y.H., Zhang, Y.W., Cang, Z.Q., et al., 2005. Environmental effects of water resource development and use in the Tarim River basin of northwestern China. Environ. Geol. 48, 202–210.

Fowler, H.J., Blenkinsop, S., Tebaldi, C., 2010. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. Int. J. Climatol. 27, 1547–1578.

Gan, T.Y., Dlamini, E.M., Biftu, G.F., 1997. Effects of model complexity and structure, data quality, and objective functions on hydrologic modeling. J. Hydrol. 192, 81–103.

Ganguly, N.D., 2011. Investigating the possible causes of climate change in India with satellite measurements. Int. J. Remote Sens. 32, 687–700.

Gao, X.J., Xu, Y., Zhao, Z.C., Pal, J.S., Giorgi, F., 2006. On the role of resolution and topography in the simulation of East Asia precipitation. Theor. Appl. Climatol. 86, 173–185.

Gleckler, P.J., Taylor, K.E., Doutriaux, C., 2008. Performance metrics for climate models. J. Geophys. Res. Atmos. 113, D06104.

Gray, L.J., Beer, J., Geller, M., Haigh, J.D., Lockwood, M., Matthes, K., et al., 2010. Solar influences on climate. Rev. Geophys. 48, 1032–1047.

Gudmundsson, L., Bremnes, J.B., Haugen, J.E., Engen, Skaugen T., 2012. Technical note: downscaling RCM precipitation to the station scale using quantile mapping - a comparison of methods. Hydrol. Earth Syst. Sci. Discuss. 16, 3383–3390.

Hirota, N., Takayabu, Y.N., 2013. Reproducibility of precipitation distribution over the tropical oceans in CMIP5 multi-climate models compared to CMIP3. Clim. Dyn. 41, 2909–2920.

Hirota, N., Takayabu, Y.N., Watanabe, M., Kimoto, M., 2011. Precipitation reproducibility over tropical oceans and its relationship to the double ITCZ problem in CMIP3 and MIROC5 climate models. J. Clim. 24, 4859–4873.

Ho, C.K., Stephenson, D.B., Collins, M., Ferro, C.A.T., Brown, S.J., 2012. Calibration strategies: a source of additional uncertainty in climate change projections. Bull. Am. Meteorol. Soc. 93, 21–26.

Huang, J.P., Guan, X.D., Ji, F., 2012. Enhanced cold-season warming in semi-arid regions. Atmos. Chem. Phys. 12, 4627–4653.

Huang, D.Q., Zhu, J., Zhang, Y.C., Huang, A.N., 2013. Uncertainties on the simulated summer precipitation over Eastern China from the CMIP5 models. J. Geophys. Res. Atmos. 118, 9035–9047.

Hunter, P., 2017. The role of biology in global climate change. EMBO Rep. 18, 673–676.

Ines, A.V.M., Hansen, J.W., 2006. Bias correction of daily GCM rainfall for crop simulation studies. Agric. For. Meteorol. 138, 44–53.
IPCC. Intergovernmental Panel on Climate Change (IPCC), 2007. Climate Change 2007: The Physical Science Basis, in the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. 2007. Cambridge Univ. Press, Cambridge, U. K., and New York.
IPCC. Intergovernmental Panel on Climate Change (IPCC), 2013. Climate Change 2013: The Physical Science Basis, in the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. 2013. Cambridge Univ. Press, Cambridge, U. K., and New York.
Ishizaki, Y., Nakaegawa, T., Takayabu, I., 2012. Validation of precipitation over Japan during 1985–2004 simulated by three regional climate models and two multi-model ensemble means. Clim. Dyn. 39, 185–206.
Kiktev, D., Sexton, D.M.H., Alexander, L., Folland, C.K., 2003. Comparison of modeled and observed trends in indices of daily climate extremes. J. Clim. 16, 3560–3571.
Kirtman, B.P.S., Adedoyin, J.A., Boer, G.J., Bojariu, R., Camilloni, I., Doblas-Reyes, F.J.F.A., Kimoto, M., Meehl, G.A., Prather, M., Sarr, A.S.C., Sutton, R., van Oldenborgh, G.J., Vecchi, G., Wang, H.J., 2013. Climate Change 2013: The Physical Science Basis. Cambridge University Press, Cambridge, UK and New York, NY.
Kovalenko, V.A., Zherebtsov, G.A., 2014. Influence of solar activity on the climate change. Atmos. Oceanic Opt. 27, 506–510.
Lafon, T., Dadson, S., Buys, G., Prudhomme, C., 2013. Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods. Int. J. Climatol. 33, 1367–1381.
Laothawornkitkul, J., Taylor, J.E., Paul, N.D., Hewitt, C.N., 2010. Biogenic volatile organic compounds in the Earth system. New Phytol. 183, 27–51.
Leander, R., Buishand, T.A., 2007. Resampling of regional climate model output for the simulation of extreme river flows. J. Hydrol. 332, 487–496.
Lenderink, G., Buishand, A., Deursen, W.V., 2007. Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach. Hydrol. Earth Syst. Sci. 11, 1145–1159.
Li, L.J., Wang, B., Zhou, T.J., 2007. Impacts of external forcing on the 20th century global warming. Chin. Sci. Bull. 52, 3148–3154.
Li, H.B., Sheffield, J., Wood, E.F., 2010. Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching. J. Geophys. Res. Atmos. 115 (-).
Li, Z., Liu, W.Z., Zhang, X.C., Zheng, F.L., 2011. Assessing the site-specific impacts of climate change on hydrology, soil erosion and crop yields in the Loess Plateau of China. Clim. Chang. 105, 223–242.
Li, Z., Zheng, F., Liu, W., Jiang, D., 2012. Spatially downscaling GCMs outputs to project changes in extreme precipitation and temperature events on the Loess Plateau of China during the 21st Century. Glob. Planet. Chang. 82-83, 65–73.
Li, C.X., Zhao, T.B., Ma, Z.G., 2014. Impacts of anthropogenic activities on climate change in arid and semiarid areas based on CMIP5 models. Chin. Sci. Bull. 59, 2972–2988 (In Chinese).
Liu, C.M., Xia, J., 2004. Water problems and hydrological research in the Yellow River and the Huai and Hai River basins of China. Hydrol. Process. 18, 2197–2210.
Ma, J.H., Wang, H.J., Ke, F., 2015. Dynamic downscaling of summer precipitation prediction over China in 1998 using WRF and CCSM4. Adv. Atmos. Sci. 32, 577–584.
Madden, R.A., Shea, D.J., 1978. Estimates of the natural variability of time-averaged temperatures over the United States. Mon. Weather Rev. 106, 1695–1703.
Mao, J.P., Robock, A., 1998. Surface air temperature simulations by AMIP general circulation models: volcanic and ENSO signals and systematic errors. J. Clim. 11, 1538–1552.
Masson, D., Knutti, R., 2011. Spatial-scale dependence of climate model performance in the CMIP3 ensemble. J. Clim. 24, 2680–2692.
Matthews, H.D., Weaver, A.J., Meissner, K.J., Gillett, N.P., Eby, M., 2004. Natural and anthropogenic climate change: incorporating historical land cover change, vegetation dynamics and the global carbon cycle. Clim. Dyn. 22, 461–479.
Maurer, E.P., Hidalgo, H.G., 2008. Utility of daily vs. monthly largescale climate data: an intercomparison of two statistical downscaling methods. Hydrol. Earth Syst. Sci. 12, 551–563.
Mcmahon, T.A., Peel, M.C., Karoly, D.J., 2015. Assessment of precipitation and temperature data from CMIP3 global climate models for hydrologic simulation. Hydrol. Earth Syst. Sci. 19, 361–377 (19,1(2015-01-21)).
Mentel, T.F., Wildt, J., Kiendler-Scharr, A., Kleist, E., Tillmann, R., Dal Maso, M., et al., 2009. Photochemical production of aerosols from real plant emissions. Atmos. Chem. Phys. 9, 4387–4406.
Miao, C.Y., Su, L., Sun, Q.H., Duan, Q.Y., 2016. A nonstationary bias-correction technique to remove bias in GCM simulations: bias-correction in the GCM simulation. J. Geophys. Res. Atmos. 121.
Milly, P.C.D., Julio, B., Malin, F., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., et al., 2008. Climate change. Stationarity is dead: whither water management? Science 319, 573–574.
Mitchell, T.D., Jones, P.D., 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. Int. J. Climatol. 25, 693–712.
Mote, P.W., Salathe, E.P., 2010. Future climate in the Pacific Northwest. Clim. Chang. 102, 29–50.
Mufti, S., Shah, G.N., 2011. Solar-geomagnetic activity influence on Earth's climate. J. Atmos. Sol. Terr. Phys. 73, 1607–1615.
Murphy, A.H., 1988. Skill scores based on the mean square error and their relationships to the correlation coefficient. Mon. Weather Rev. 116, 2417–2424.
Narisma, G.T., Foley, J.A., Licker, R., Ramankutty, N., 2007. Abrupt changes in rainfall during the twentieth century. Geophys. Res. Lett. 34, 306–316.
Panofsky, H.A., Brier, G.W., 1968. ome Application of Statistics to Meteorology, Earth and Mineral Sciences Continuing Education. College of Earth and Mineral Sciences.
Park, C., Min, S.K., Lee, D., Cha, D.H., Suh, M.S., Kang, H.S., et al., 2016. Evaluation of multiple regional climate models for summer climate extremes over East Asia. Clim. Dyn. 46, 1–18.
Phillips, T.J., Gleckler, P.J., 2006. Evaluation of continental precipitation in 20th century climate simulations: the utility of multimodel statistics. Water Resour. Res. 42, 446–455.
Piani, C., Haerter, J.O., Coppola, E., 2010. Statistical bias correction for daily precipitation in regional climate models over Europe. Theor. Appl. Climatol. 99, 187–192.
Pierce, D.W., Barnett, T.P., Santer, B.D., Gleckler, P.J., Thiemens, M.H., 2009. Selecting global climate models for regional climate change studies. Proc. Natl. Acad. Sci. U. S. A. 106, 8441–8446.
Pitman, A.J., 2010. The evolution of, and revolution in, land surface schemes designed for climate models. Int. J. Climatol. 23, 479–510.
Reynolds, J.F., Smith, D.M., Lambin, E.F., Nd, T.B., Mortimore, M., Batterbury, S.P., et al., 2007. Global desertification: building a science for dryland development. Science 316, 847–851.
Richardson, A.D., T Andy, B., Philippe, C., Nicolas, D., Friedl, M.A., Nadine, G., et al., 2010. Influence of spring and autumn phenological transitions on forest ecosystem productivity. Philos. Trans. R. Soc. Lond. 365, 3227–3246.
Richardson, A.D., Keenan, T.F., Migliavacca, M., Ryu, Y., Sonnentag, O., Toomey, M., 2013. Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. Agric. For. Meteorol. 169, 156–173.
Rong, H.Y., Shreedhar, M., Stefan, U., 2013. Downscaling daily precipitation over the Yellow River source region in; China: a comparison of three statistical downscaling methods. Theor. Appl. Climatol. 112, 447–460.
Rotenberg, E., Dan, Y., 2010. Contribution of semi-arid forests to the climate system. Science 327, 451–454.
Schmidli, J., Frei, C., Vidale, P.L., 2006. Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods. Int. J. Climatol. 26, 679–689.
Schuenemann, K.C., Cassano, J.J., 2009. Changes in synoptic weather patterns and Greenland precipitation in the 20th and 21st centuries: 1. Evaluation of late 20th century simulations from IPCC models. J Geophys. Res. 114, D20113.
Shindell, D.T., Schmidt, G.A., Miller, R.L., Mann, M.E., 2003. Volcanic and solar forcing of climate change during the preindustrial era. J. Clim. 16, 4094–4107.
Sillmann, J., Kharin, V.V., Zhang, X., Zwiers, F.W., Bronaugh, D., 2013. Climate extremes indices in the CMIP5 multimodel ensemble: part 1. Model evaluation in the present climate. J. Geophys. Res. Atmos. 118, 1716–1733.
Song, F.F., Zhou, T.J., 2014. Interannual variability of east Asian summer monsoon simulated by CMIP3 and CMIP5 AGCMs: skill dependence on Indian Ocean-Western Pacific anticyclone teleconnection. J. Clim. 27, 1679–1697.
Su, F.G., Duan, X.L., Chen, D.L., Hao, Z.C., Cuo, L., 2013. Evaluation of the global climate models in the CMIP5 over the Tibetan Plateau. J. Clim. 26, 3187–3208.
Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res. Atmos. 106, 7183–7192.
Teutschbein, C., Seibert, J., 2012. Bias correction of regional climate model simulations for hydrological climate-change impact studies: review and evaluation of different methods. J. Hydrol. 456-457, 12–29.
Themeßl, M.J., Gobiet, A., Leuprecht, A., 2011. Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. Int. J. Climatol. 31, 1530–1544.
Thevs, N., Peng, H., Rozi, A., Zerbe, S., Abdusalih, N., 2015. Water allocation and water consumption of irrigated agriculture and natural vegetation in the Aksu-Tarim river basin, Xinjiang, China. J. Arid Environ. 112, 87–97.
Van Roosmalen, L., Sonnenborg, T.O., Jensen, K.H., Christensen, J.H., 2011. Comparison of hydrological simulations of climate change using perturbation of observations and distribution-based scaling. Vadose Zone J. 10, 136–150.
Walsh, J.E., Chapman, W.L., Romanovsky, V., Christensen, J.H., Stendel, M., 2008. Global climate model performance over Alaska and Greenland. J. Clim. 21, 6156–6174.
Wang, L., Chen, W., 2014. A CMIP5 multimodel projection of future temperature, precipitation, and climatological drought in China. Int. J. Climatol. 34, 2059–2078.
Wilby, R.L., Wigley, T.M.L., Conway, D., Jones, P.D., Hewitson, B.C., Main, J., et al., 1998. Statistical downscaling of general circulation model output: a comparison of methods. Water Resour. Res. 34, 2995–3008.
Wu, J., Gao, X.J., 2013. A gridded daily observation dataset over China region and comparison with the other datasets. Chin. J. Geophys. 56, 1102–1111 (in Chinese).
Xu, Y., Xu, C.H., 2012. Preliminary assessment of simulations of climate changes over China by CMIP5 multi-models. Atmos. Ocean. Sci. Lett. 5, 489–494.
Yue, T.X., Zhao, N., Fan, Z.M., Li, J., Chen, C.F., Lu, Y.M., et al., 2016. CMIP5 downscaling and its uncertainty in China. Glob. Planet. Chang. 146, 30–37.
Zhao, Y., Yu, X.X., 2013. Effects of climatic variability and human activity on runoff in the Loess Plateau of China. For. Chron. 89, 153–161.
Zhao, X.C., Wu, H.B., He, L., 2008. Estimate of climatic noise and potential predictability of seasonal precipitation in China. J. Nanjing Inst. Meteorol. 6, 819–827 (In Chinese).
Zhao, T.B., Chen, L., Ma, Z.G., 2014. Simulation of historical and projected climate change in arid and semiarid areas by CMIP5 models. Chin. Sci. Bull. 59, 412–429.
Zhao, T.B., Li, C.X., Zuo, Z.Y., 2016. Contributions of anthropogenic and external natural forcings to climate changes over China based on CMIP5 model simulations. Sci. China Earth Sci. 59, 503–517.
Zherebtsov, G.A., Kovalenko, V.A., Kirichenko, K.E., 2015. Main physical processes and mechanisms responsible for the observable climate changes in the 20–21st centuries. International Symposium Atmospheric & Ocean Optics: Atmospheric Physics.
Zhou, T., Li, Z., 2002. Simulation of the east Asian summer monsoon using a variable resolution atmospheric GCM. Clim. Dyn. 19, 167–180.